Journal of
Experimental Psychology

VOL. 20, NO. 2 FEBRUARY, 1937

RESPONSE IN THE MEDIAN PLANE LOCALIZATION OF SOUND *

BY C. H. PEARCE

In a previous communication the writer reported [1] indications that a particular response situation might play a part in the alteration of the nature of the localization of a source of sound. The situation was one in which the observer responded to sounds originating in the median plane by inserting a stylus into the holes of a board. In that experiment it was sometimes noticed that if the response was toward the left of the board the sound was heard not in the median plane, but toward the left. Similarly, median plane sounds often appeared to come from the right when the subject responded toward the right of the board. The phenomenon occurred with several subjects of a group who were well aware that the sounds were all situated in the median plane, but who still reported that the sounds appeared definitely lateral. The purpose of the present investigation was to determine whether this phenomenon would appear with regularity if the subject did not know the exact source of the sound.

EXPERIMENTAL SITUATION

The apparatus used is shown in Fig. 1. The sound sources were five phones arranged in the median plane above the subject's head on a metal arc, and separated

* From the Psychological laboratory of Princeton University.

[1] C. H. Pearce, Training in median plane localization of sound, *J. Gen. Psychol.* (in press).

by equal spatial intervals. The clicks at the five positions were qualitatively distinguishable. The phones were actuated by the current from a dry cell, which was discharged into the phone after being shunted from a condenser with a capacity of .5 microfarad. A series of switches, one to each phone, was operated from a position in front of the subject. The subject was blindfolded and never saw the apparatus. The head was securely held by a device which exerted tension at the forehead and occiput.

FIG. 1.

A semicircular board in front of the subject was used by him to indicate the location of the sound. The board will be described below. The experiment was conducted in a sound-proof room.

Twenty male undergraduate students of psychology served as subjects. Of these six constituted what will be termed a control group. The remainder constituted two groups of seven subjects each, which groups will be known as experimental groups I and II. Each subject served for four or five sessions. Each session lasted from thirty to fifty-five minutes. The average number of judgments for a subject was near 800, or about 160 for each stimulus position.

The instructions to the subjects of the control group follow. "This is a problem in auditory localization. You will hear clicks in various positions around the head, and you are to tell where these sounds come from. If you will think of the region around the head as an inverted bowl or umbrella over your head you will have an idea of the locus of the stimuli. It is obvious that there are two dimensions here, and you will have to locate the sound with respect to both. The dimensions are front and back, and left and right. If we consider the front-back dimension there are five points in the median plane, stretching from front to back above the head around which the sounds may appear. The middle of these positions is straight up. The lowest in front is just above the level of the eyes, while the lowest in back is just above the level of the ears. The other two positions are half way between straight up and the other two positions. These positions are numbered from 1 to 5, starting in front. After each presentation you will tell which number you think most nearly fits the position of the sound. However, and this is important, clicks may appear in the left-right dimension as well as in the median plane at any of these five positions. So after you give the position number say whether the sound is left, middle or right. After each response I shall tell you the correct number of the position. Try to use my corrections of the distance front or back to improve your localizations. I shall not correct you on the report as to left, middle or right. You will have to do the best you can in these reports. The purpose of the experiment is to see whether as you improve in the front-back judgments with training you will also improve in the localization as to left and right without training."

The subjects of the experimental groups received the same instructions, save that instead of using numbers they indicated the anterior-posterior positions by inserting the stylus into a particular hole in the board. With respect to the mid line of the subject's body these holes were placed two to the left, two to the right, and one in the middle. The holes were equally spaced about the edge of the board. Each hole was used to indicate a particular position in front or in back. If we number these positions from 1 to 5 as in the control group the hole furthest left was used for position 1 while the hole furthest right stood for position 5 in the first experimental group. The subjects of group II used just the reverse order, starting at the right with position 1 and ending at the left with position 5. The experimenter corrected each manual response by withdrawing the inserted stylus and pushing it into the correct hole. The subject identified this hole and withdrew the stylus. The verbal judgment as to whether the sound was lateral was not corrected. Between stimulus presentations the stylus was held in the subject's hand. No directions as to which hand he was to use when inserting the stylus were given. The subjects were all right handed as far as the grasping of the stylus was concerned. The left hand was observed to be infrequently used. When it was used the stylus was usually being held by both hands.

In a preliminary series it was found that certain subjects responded to one side or the other almost exclusively for the reason that they located the sounds in large measure in front or in back. It was necessary for the experiment that the subjects be corrected for the purpose of obtaining responses to both left and right.

With the experimental situation stated it becomes possible to explain the relationships which were to be tested. From

the results of the previous investigation it was expected that in experimental group I, where a response to the left indicated a sound in front, clicks in positions 1 and 2 would be more frequently judged to be left than right, while clicks in positions 4 and 5 should be more often perceived as right than as left. The second experimental group was expected to show a reverse relationship between stimulus position and lateral displacement; namely, sounds at positions 1 and 2 should appear preponderantly right while sounds at positions 4 and 5 should appear more often as left. The control group was to be used to compare the percentages of lateral displacements in subjects making no manual response with the same percentages in the experimental groups.

Results

Figure 2 shows the relative percentages of responses distributed to left and middle and right by reference to the accompanying verbal report. The percentages for each posi-

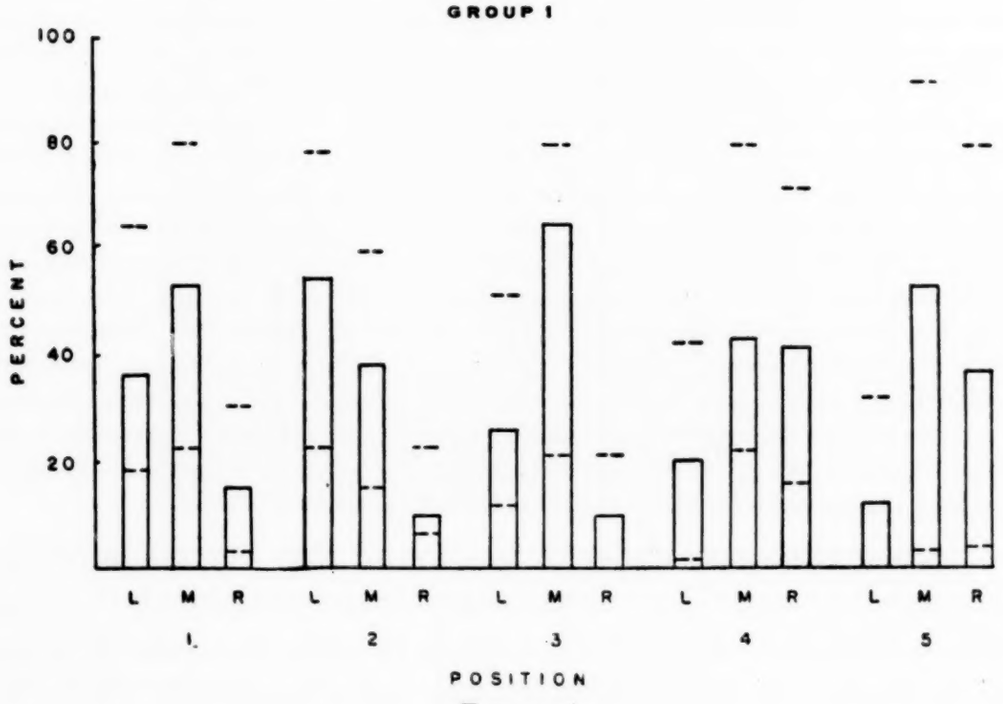

Fig. 2.

tion represent the average for the seven subjects of experimental group I. The dotted lines accompanying each bar indicate the range for that group. The positions are those

selected by the subjects, without reference to the actual stimulus position. The diagram shows that for group I report more clicks from the left when they are responding to the left than they report clicks to the right, and conversely report more clicks to the right when responding to the right.

Figure 3 presents the same data as Fig. 2, but for group II.

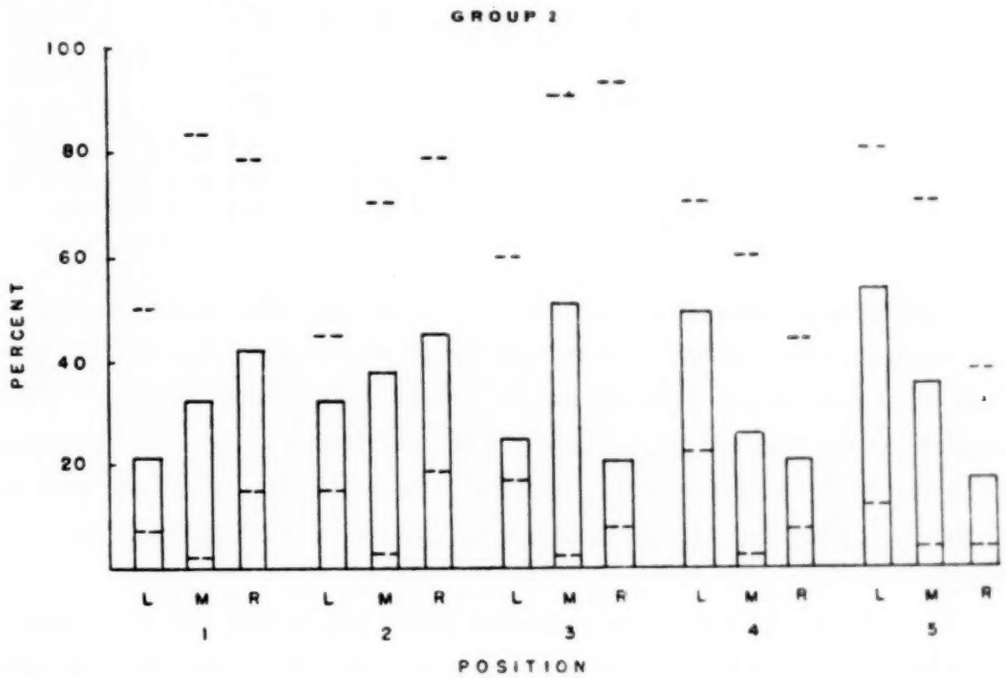

FIG. 3.

Here the same generalization as that made for group I holds true, so that the relation observed between report of laterality and direction of response holds even if the same stimulus evokes movement to different sides, as it does in the two groups.

Of the six control subjects two reported no lateral clicks, while one reported that all clicks appeared to come from a constant short distance to the left. The results of the three remaining subjects are not sufficiently similar to allow them to be likened to those of either experimental group. Table I presents the results for the subjects of the control group according to the analysis of Figs. 2 and 3.

In Table I it will be noted that while there is little change in the ratio of the percentages of left and right judgments as passes from position to position 5 there is a slight preponder-

TABLE I

PERCENTAGES OF SOUNDS PERCEIVED AS LEFT, MIDDLE OR RIGHT BY SUBJECTS OF THE CONTROL GROUP

Subject	Stimulus Positions														
	1			2			3			4			5		
	L	M	R	L	M	R	L	M	R	L	M	R	L	M	R
1	0	100	0	0	100	0	0	100	0	0	100	0	0	100	0
2	0	100	0	0	100	0	0	100	0	0	100	0	0	100	0
3	100	0	0	100	0	0	100	0	0	100	0	0	100	0	0
4	8	91	1	9	86	5	9	90	1	11	83	6	8	90	2
5	34	23	43	42	42	16	28	45	27	26	48	26	26	40	34
6	7	87	6	23	59	18	32	50	18	34	50	16	24	76	0

ance of placements to the left. There are also indications in the results shown for the subjects of the experimental groups that sounds were heard more often left than right. This tendency suggests that either one or several of the sources were actually slightly left, or that the left side was favored to a small degree by the reflecting surfaces of the apparatus and room.

The data thus far considered have included all responses, whether or not the responses were correct with respect to the position of the stimulus in the mid line. The next two figures, 4 and 5, include only reports of left or right in which there were no reversals from front to back or vice versa. This means that if the stimulus was at position 4 the judgment of right or left is included in this analysis only if the response was to positions 4 or 5, while if the stimulus was at position 1 the judgment of right or left is included only if the response was to positions 1 or 2. All cases in which stimulus position or response position 3 occurred are omitted from these figures. The height of each bar shows the percentage of the cases defined above in which stimuli at positions 1 and 2 were called left and in positions 1 or 2, and stimuli at positions 4 and 5 were called right and in positions 4 and 5. Conversely, the difference between the height of any bar and the 100 percent line represents the percentage of responses in which sounds at positions 1 and 2 were called right and in positions

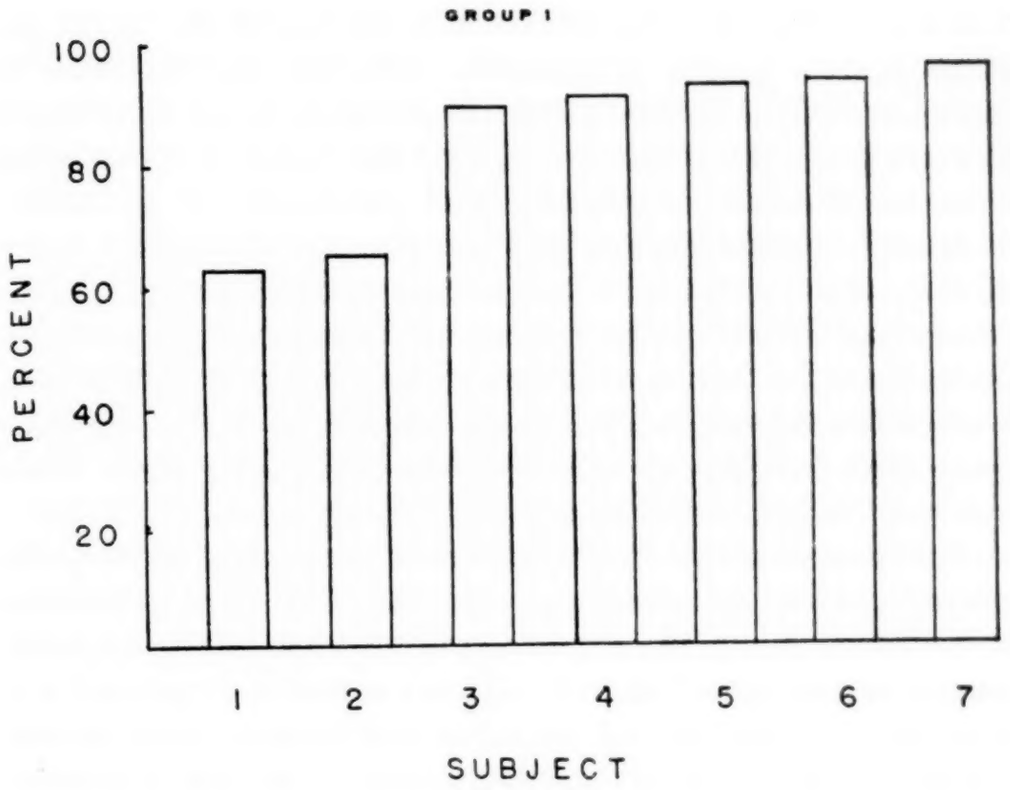
GROUP 1
100
80
60
40
20
PERCENT
1
2
3
4
5
6
7
SUBJECT

FIG. 4.

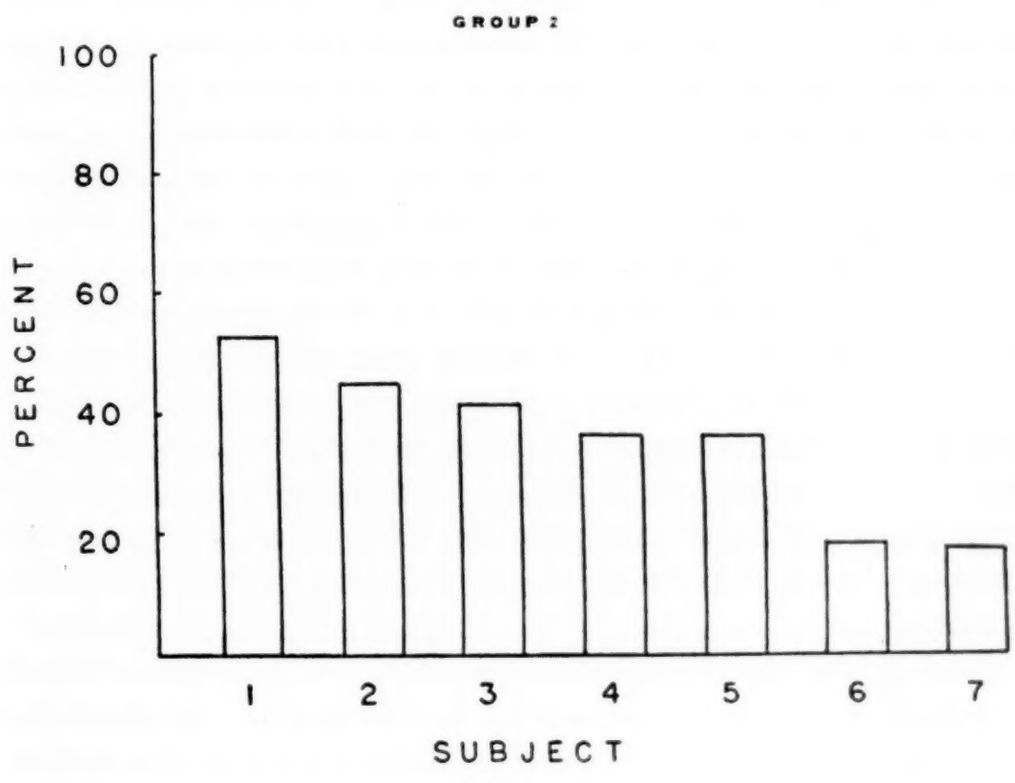
GROUP 2
100
80
60
40
20
PERCENT
1
2
3
4
5
6
7
SUBJECT

FIG. 5.

1 and 2, and so on. In other words, the higher the bar is, the more perfect is the relationship between the direction of response and the side to which the sound is heard, in group I. For group II the lower the bar is, the closer is the relation between direction of response and perception of the sound as lateral. Consideration of Figs. 4 and 5 shows that seven of the fourteen subjects of the experimental groups gave a percentage of about 84 percent or better in this analysis. These data are not as reliable as those of the previous summary, since the subjects who were not precise in their localizations with respect to the numbered positions have fewer results included in this analysis.

Little systematic introspection was required during the progress of the experiments. The subjects were questioned at the conclusion of the experimental series as to the apparent sources of the sound stimuli. These questions together with voluntary comments and recorded localizations provide considerable material for the determination of the subjective appearance of the sounds.

Of the twenty subjects three stated some doubt as to believing that all the sounds were where the experimenter had said they might be. Subject 1 of the control group who reported no sounds to left or right asked whether there were any sounds actually coming from the right or left. Subject 5 of group I stated that all the sounds appeared to come from the front in position 1. He actually responded as though there were three or four positions and chose position 2 in the majority of cases. He concluded that there was but one source, situated in position 1, and that the reason some sounds appeared to be coming from the left or right was because he was responding to left or right. A third subject, number 6 of group I, reported in his first sitting that he felt as though he were being influenced in his lateral placements by the board.

As previously indicated, subjects did not place equal numbers of sounds at each of the five positions. Some of the observers tended to group their judgments about one or two positions so that a majority of localizations were to the front

or rear. Some of the subjects located all sounds in front while others located all of the sounds in the rear. Since the sounds at different positions were qualitatively distinguishable subjects soon found that the experimenter's corrections indicated which sounds were actually in the rear and which were really in front. The attempt of the subject to improve in his performance caused a condition in which a sound apparently in the rear might come to be identified as a sound actually coming from the front, and be so indicated by the response, or in the control group, the reported position. In view of these circumstances it is to be expected that in the experimental groups different arrangements of the five sounds on the left and right would be reported by different subjects according to the locality of the majority of the localizations with respect to the mid line.

Five of the fourteen subjects of the experimental groups placed a large majority of the clicks in front. Two of these were in group I and three were in group II. The two subjects of group I reported at the conclusion of the experiment that the five sounds were situated horizontally at the level of the eyes, with sounds 1 and 2 on the left and sounds 4 and 5 on the right, with position 3 varying about the middle in the same line. The three observers of group II reported a similar spreading out in front of the head with, however, positions 1 and 2 on the right and positions 4 and 5 on the left. Similar but reversed relations held in general for three subjects who placed most of the sounds to the rear. These observers, all in group II, reported the positions to be spread out in the rear from right to left, starting with position 1. The remainder of the subjects of the experimental groups localized the sounds approximately where they were with respect to front and rear, although the positions did not usually appear equally apart spatially. The general impression of these subjects was that the sounds formed a diagonal line from front to back, starting in front on the left or right as the case might be, proceeding through the center directly above the head and passing to right or left.

In the control group two made most localizations in front

while one heard most of the sounds in back. The other three subjects gave a more random distribution of choices. Not one of these subjects finally reported any definite lateral arrangement of the sounds, save subject 2 who reported all sounds as an equal and short distance to the left.

The apparent lines into which the five sounds fell as presented above were not usually strict repetitions of the spatial arrangement of the stimuli from front to back. For a few subjects sounds in positions 2 and 4 would appear more distant laterally than those in positions 1 or 5; or positions 2 and 1 might appear to be very close together or identical in position, and similarly with positions 4 and 5.

From these results it is impossible to answer the obvious question of whether a more extensive movement to one side, such as reaching to position 1 on the left as compared with reaching to position 2 would, after a series of judgments, cause the sound at position 1 to be perceived as further left than that at position 2. To investigate the possibility of obtaining such results it would be necessary to employ greater separation of the board positions and also to secure a more substantial number of choices of the extreme positions than are available here.

Another point at which the results are not uniform from subject to subject concerns the course of the change in the number of lateral judgments from sitting to sitting. The average percentage of lateral placements for the fourteen subjects of the experimental groups is 56 percent in the first sitting and 66 percent for the fifth sitting. Of these subjects two showed a steady reduction of placements to the sides in successive sittings amounting to a total of 59 percent in one case and 38 percent in the second. Of the remaining subjects nine increased lateral placements by percentages ranging from 15 percent to 35 percent in five sittings while three varied less than 10 percent in successive sittings.

A similar degree of variation to that just mentioned is discovered in the relations of the direction of response to the position of the stimulus in succeeding sessions. While the majority of subjects showed an increasing tendency to place

sounds to which they responded manually to the left on the left, etc., several maintained percentages which varied little throughout the series and two subjects showed a reverse tendency to that of the majority. It is possible that these variations are due to differences in the attitudes of the subjects but a positive conclusion on this point cannot be made on the basis of the present data.

Discussion and Conclusions

It is concluded that the results in the present situation show some relation to hold between the localization of sounds presented in the median plane and the response made to those sounds. When the anterior-posterior position of a sound was indicated by a movement to the right that sound was perceived more often as right than as left. If the same position was indicated by a movement of the arm to the left that sound was more often perceived as left than as right.

It is not strange that sounds originating in the median plane should be sometimes reported as lateral. This has been found to happen even when the subject is well aware that the source is in the median plane. The consideration that the head cannot be expected to remain absolutely immobile for thirty minutes is sufficiently suggestive for the present situation in view of the extremely small movement of the head necessary to alter binaural stimulus ratios. There is, however, no reason to suppose that chance variations of head position would cause the relation of displacements to position of stimulus which holds in this experiment. Moreover, the very fact that the subjects of the experimental groups reported many more lateral sounds than did the subjects of the control group shows that head movement, which is a possible factor in all three groups of subjects, does not explain the observed relations.

Two of the subjects of the experimental groups, as has already been noted, were suspicious of the board as an influence in their reports of laterality. It is significant that both of these observers showed a very marked positive relationship between the direction of the response and the

side to which particular stimuli were heard. It might be expected rather that such suspicions would operate in an opposite manner. One of these subjects was very critical about the situation. He asked many questions and decided that there was but one sound source and a single position. The other subject reported that he did not reflect about the situation further than his comment about the influence of the response board. These subjects are mentioned because they represent individual variations which are found more or less suggestive for all the subjects.

Beyond the considerations which may be supposed to reflect individual variation in any experiment on sound localization there are two factors in the present situation which should be referred to explicitly. Individual criteria for what constituted the median plane were broader or narrower from subject to subject. This means that sounds which would be judged as very near the median plane but lateral by one subject would be judged as actually median by another subject or even by the same subject at a different time. This circumstance explains partly the individual variations in the number of sounds reported as lateral. Another factor has to deal with the fact that the subject had two judgments to make. One of these judgments, indicating the position of the sound in the dimension stretching from front to rear, was subject to correction. The other was not. Since the subject was trying to improve his percentage of correct manual responses it is natural that the corrected aspect of the response situation would be favored with more attention or effort than would be the judgments as to laterality. It could be observed that some subjects would make the first, manual response and then fail to make the verbal report, stating on questioning that they had not noticed whether the sound was middle or lateral because they had been preoccupied with the first report. When the results of the present experiment are considered such considerations as these should not be neglected.

The results of this investigation are interpreted to mean that in the absence of visual cues auditory 'thereness' is

shown to be in part a function of movement of gross body parts. This interpretation supports the hypothesis that auditory localization is due in some part to a change in the position of body segments with reference to a sound source. In the normal course of events such responses become movements *toward* a source of sound because cues afforded by vision touch and changes in auditory stimulation provide the basis for a correct orientation. In the present situation movement is not towards the sound, as in adaptive auditory localization, and this fact is reflected in the apparent displacement of the sound source from its 'true' position. Here binaural stimulus conditions are not as effective as they would be in situations where a more specific response occurs. This fact raises the question of the degree to which the response factor and the binaural stimulus conditions interact in ordinary experience. That response can at all modify the effect of purely auditory cues is a consideration which must be taken into account in a theory of auditory space.

Summary

1. Fourteen subjects were observed to perceive sounds originating in the median plane as though the stimulus positions were sometimes to the left and right.

2. The apparent location of the sounds was found to be variable with the mode of response so that sound stimulation from a point in the median plane could be perceived as left or right in a majority of cases according as the sound evoked a manual response to the left or right.

3. Six subjects who responded verbally to the same sounds in the median plane either failed to perceive the sounds as lateral in position or reported lateral displacements not consistently related to each other or to the spatial arrangement of the sound stimuli.

4. The perception of the sounds was observed to be related to the original localizations of the sounds with respect to front and rear; *i.e.*, sounds originally heard in back and in the median plane came to be heard laterally but still in the rear.

REVERSAL OF AUDITORY LOCALIZATION *

BY CLARENCE F. WILLEY, EDWARD INGLIS AND C. H. PEARCE

The experiment of P. T. Young on the perceptual changes occurring with acoustical transposition of the ears has attracted much attention by reason of the implications of his results for empiristic theories of localization. His work, like that of Stratton in vision, has been extensively cited as evidence for the possibility of a very considerable readjustment of localization habits within a comparatively short time. Young (**5**, 413) reported that although "from first to last purely auditory localizations were reversed," . . . "normal localization usually occurred when the source was fixated visually or when an attentive adjustment was made to its position." He stated further that as the experiment progressed "a casual glance at an object was sufficient to check up the fact that its sound was normally localized. With habituation, all the sounds in a complex situation were normally localized without any thought about the matter." The sharp distinction made by Young between *auditory* localization and *visual-auditory* localization is apparent in the foregoing quotations. The former type of localization remained unchanged during the experimental period, while the other "developed so that the sight-sound coördinations were indistinguishable from normal life" (**5**, 424).

In view of the great theoretical significance which has been attached to these observations, the present authors believed that an experiment should be performed in which more than one subject would wear a similar pseudophone over a longer continuous period of time. In this experiment two observers wore pseudophones for seven and eight days respectively, each totalling approximately 100 hours. A third observer wore a pseudophone for three days.[1]

* From the Psychological Laboratory of Princeton University.

[1] Three days is the length of time Young wore his pseudophone continuously. Removal of the pseudophone was necessitated by irritation of the sensitive tissues in contact with the ear pieces.

The conditions of the experiment were substantially the same as those of Young's investigation. The pseudophones used did not have the large receiving-trumpets which were such a conspicuous feature of Young's apparatus, since it was not considered that these increased the efficiency of the device. The tubes which transmitted the sound were made of ¼ inch copper tubing. They were covered with thick-walled rubber tubing to provide additional insulation. Hard rubber ear pieces were attached to one end of each tube; the other end was bent out in such a way that the external aperture was normal to the binaural axis. The ear-pieces were made oval in shape to fit the contour of the auditory canal. They were covered with short sections of soft rubber tubing in order to reduce friction caused by movements of the head and lower jaw. Vaseline and cotton were packed around the ear-pieces to provide better sealing. Both tubes were rigidly attached to a single strip of resilient metal which lay on the top of the head and provided sufficient tension to hold the ear-pieces in place. In spite of the precautions taken to minimize irritation, the constant pressure exerted by the ear-pieces caused considerable discomfort toward the end of the experiment.

The tubes (including the ear-pieces) were slightly over 13 inches in length. There was considerable resonance for low tones under certain acoustic conditions. This was particularly noticeable in a large dining hall where the lower frequencies made understanding of speech somewhat difficult. Outdoors, and in rooms of ordinary size, the effect of resonance was barely appreciable.

The pseudophones were worn throughout the day. They were removed at night, and the ears were then packed with vaseline and cotton. Each observer kept a record of his auditory experiences during the experimental period. These were compared from time to time to determine how generally a given phenomenon was observed. For the most part the reports of the three observers agreed very well. Many of the observations were made independently by two or all three of the observers. Discussion will be confined to those results which appeared repeatedly or which could be confirmed by the reports of all the experimenters.

Tests of acuity were made under different conditions by means of a Western Electric 2-A Audiometer. The hearing loss for each ear was tested for each of the following frequencies: 128, 512, 1024, 2048 and 8192 cycles. The head was held in a fixed position and the ear-phone was placed at a distance of three inches from the ear or the open end of a pseudophone tube. In Series I of the tests the acuity was tested without the pseudophones. In Series II each ear was tested with the pseudophone on and both pseudophone tubes open. In Series III each ear was tested with the open end of the tube nearest the ear-phone covered with a hard rubber cap; the other tube was left open. In Series IV, the acuity of each ear was tested with both of the tubes capped. In Series I, III and IV, sounds from the ear-phone were localized on the same side as the ear-phone. In Series II sounds from the ear-phone were localized on the right when the ear-phone itself was on the left, and *vice versa*. A comparison was made of the results of Series I for one ear and Series II for the opposite ear to obtain a measure of the loss of acuity which accompanied the wearing of the pseudophone. For 1024 cycles this loss was 10 to 15 decibels. A comparison of the results of Series II and Series III provided a measure of the leakage of sound through the cotton and vaseline insulation. The smallest difference obtained between a threshold of Series II for one ear and the corresponding threshold of Series III for the other ear for any frequency was 10 decibels. In the majority of determinations a larger difference was obtained. For 1024 cycles the range was 10 to 20 decibels. These figures represent the difference in intensity between the sound carried through the uncapped tube to the ear on the side opposite the audiometer phone and the sound reaching the ear on the same side as the phone. The sound reaching the ear on the same side as the ear-phone in Series III may reach that ear either by leakage through the insulating vaseline and cotton or by passing around the head and returning through the tube which is uncapped. The difference between the thresholds of Series III and Series IV throws light on the relative importance of these two paths. With both tubes

capped, as in Series IV, an additional hearing loss was sometimes obtained, but in most of the comparisons there is no difference between the thresholds of Series II and IV. Thus, although there is sometimes a slight decrease in acuity resulting from the capping of the tube opposite the ear-phone, it is evident that leakage through the packing on the same side as the ear-phone was primarily responsible for the values obtained in Series III.

It appeared from these results that the conditions of auditory stimulation in an experiment of this sort are by no means a simple reversal of the conditions of normal hearing. On the contrary one is dealing with exceedingly complex stimulation. The pattern of sound impinging on the ear-drums differs greatly from the original sound. A sound coming from the left side, for example, gives rise to auditory stimulation in each of the following ways: (1) by passing through the pseudophone tube to the right ear; (2) by passing around the head and leaking through the cotton and vaseline insulation to the right ear; (3) by leaking through the cotton and vaseline insulation to the left ear, and (4) by passing around the head and returning to the left ear through its pseudophone tube. The sound reaching the right ear through its tube is at least ten decibels louder than that which reaches either ear in any other way, and so may be expected to be interpreted as the 'near' sound on the basis of intensity discrimination. Sounds reaching the right ear by (1) and (2) probably differ little except in intensity. The phase relationships for the right ear would not be appreciably different since the distance around the head through the tube is approximately the same as outside the tube. On the other hand, the sounds reaching the left ear by (3) and (4), respectively, will be out of phase by an amount determined by the time required for a sound to traverse a distance corresponding roughly to twice the diameter of the head. The sound leaking through the insulation to the left ear from a source at the left will lead the sound reaching the left ear through the tube by approximately that distance.

Numerous other complicating factors are also operative.

Some leakage must take place over the entire length of the tube. The distance between the pinnæ of the pseudophone is over 20 percent greater than that between the natural pinnæ. In the light of all these considerations it becomes evident that an observer in an experiment of this sort is receiving a complex of stimuli only imperfectly resembling a simple transposition of right and left auditory stimulation. The effect of the pseudophone can hardly be regarded as equivalent to that which would be produced "If a master surgeon could transplant the right inner ear to the left side of the head . . . without disturbing any of the neural connections, and at the same time transplant the left inner ear to the right side of the head" (**5**, 399). Young's belief that the pseudophone produces the same result as such an operation is justified only if it is possible to accept his assumption that "by far the greater part of the energy reaching the ears came through the apparatus" (**5**, 401). It was found in the present experiment that the ears were receiving normal stimulation through leakage only 10 to 20 decibels less intense for tones of 1024 cycles than the reversed stimulation passing through the tubes. The objection may be raised that the pseudophone used in the present experiment was less efficient than that of Young. A direct comparison of the efficiency of Young's pseudophone with that of the type used in the present experiment is impossible as Young made no quantitative estimate of leakage. It is probable, however, that the efficiency of the two devices was approximately the same, since the watch test for leakage which Young employed gave the same results when used in this experiment. When a satisfactory seal was secured by packing vaseline and cotton around the ear pieces, the sound of a watch was clearly reversed when the watch was held opposite the open end of a tube and was not perceptible when the watch was moved away from the opening closer to the ear on the same side.

The pseudophone is clearly not a perfect instrument for the transposition of auditory stimulation. In view of the complexity of the stimulation received by the ear it is perhaps surprising that in the present experiment the reversed locali-

zation was so definite and so easily dissociated from the visual representation of the sound-producing object.

Tests of localization were made at intervals during the experiment. In order to eliminate the influence of reverberation, these tests (as well as the acuity measurements) were made in a room with Celotex walls. One method of testing localization consisted in securing judgments of the apparent position of different sounds. Both a metronome and an oscillator and speaker were used as the sound-producing devices. The metronome or speaker could be moved in a circle of 50 cm radius in the transverse plane at the level of the ears. It was found that the most accurate lateral reversals were obtained when the object producing the sound was gradually moved out from the median plane. In the procedure used the source of sound was moved out from a position either directly in front of or directly behind the head. The observer was instructed to make a judgment of 'right' or 'left' as soon as the localization seemed unequivocal. Judgments of this type which had been made during the preliminary experimentation before the pseudophones were worn continuously permitted comparisons with the conditions of normal hearing. When the observers were not wearing pseudophones, correct judgments were almost invariably obtained before the metronome had been moved out 3½ inches from the median plane. The angular displacement necessary to secure correct localization varied from three to ten degrees. Delaying the report until the observer felt certain of his judgment possibly made these values higher than they otherwise would have been. With the pseudophones, the source of sound had to be moved out approximately twice as far to secure unequivocal judgments of localization. The results obtained with use of the speaker (with tones of 128 and 1024 cycles) and the metronome were substantially the same in this respect. Out of 200 'right' and 'left' judgments of this sort made while the subjects were wearing pseudophones, 92 percent were reversed.

Under the conditions of normal hearing all three subjects showed a tendency to give a preponderance of 'front' judg-

ments for an equal number of front and back presentations of the stimuli. With the pseudophones the opposite tendency was noted.

The experimenters pursued their regular activities during the experiment. Casual observations proved to be a valuable source of information regarding the characteristics of pseudophonic perception. It was noted throughout the experiment that human speech under certain conditions gave particularly definite reversal. This was especially noticeable when two people whom the observer regarded peripherally were conversing on opposite sides of the observer. Reversal occurred whether they were speaking alternately or simultaneously. In other situations, which will be discussed below, speech was normally localized.

Reversal of all the ordinary street sounds was noted by all the subjects from the first day of the experiment. The barking of dogs, the sound of a starting train, and the exhaust of passing automobiles were heard on the side opposite their true positions. The noise from moving automobiles was not only subject to lateral reversal but to front-back and back-front reversal as well. These 180 degree reversals were obtained when the automobile was either overtaking or meeting the observer; they were equally distinct when the vehicle was in or out of the field of vision.

The sound of an airplane in the air generally was not experienced as a complete 180 degree reversal, although lateral shifts of localization were regularly observed. One subject, however, reported an instance of a diametrical reversal of the sound of an airplane; on this occasion the characteristic sound seemed to be coming out of the ground.

Perhaps the most striking reversals obtained were those of musical sounds. Shifts in the localization of music were sometimes astonishingly clear; they were not usually affected by knowledge of the source. Pipe organ and choral music produced typically compelling reversals. Music of the former type appeared to be issuing from the echo pipes at the opposite end of the auditorium, except during the passages in which those pipes were actually being used. The radio was found

to be a convenient and effective means of testing reversals. Front-back and back-front reversals were frequently reported; the former were obtained when the observer was standing directly in front of the radio and facing it. For all positions front-back and back-front reversals were at least as generally observed as simple lateral reversal. This held true for different radios and for rooms of different dimensions. Lateral reversals of radio sounds were not appreciably influenced by the obvious absurdity of the apparent position of the source of sound. Music would be heard as issuing from a fireplace, a closet or a bare wall. A radio in a room on the second floor of a dormitory was localized by two of the observers who were on the ground outside the building as above them in the air on their opposite side and level with the window.

Numerous incorrect motor adjustments were made, particularly at the dinner table, where the complexity of the sound stimulation was especially favorable to confused localization. At first each observer found himself invariably turning the wrong way when addressed by the waiter or by persons sitting on either side. After the pseudophones had been worn for a time, a conscious correction of movements became the rule when it was necessary to utilize auditory cues exclusively. Whenever the observer was startled the immediate motor response was always to the wrong side.

It is a familiar fact that deliberate searching movements increase the exactness of localization under the conditions of normal hearing. Young states that he sometimes found the same to be true of pseudophonic perception. Out of 125 trials in which 'auditory-motor' search was the basis for judgments of the position of a clock rotated about his head, he succeeded in making correct judgments in 96 percent of the cases. In about half of these there was a discrepancy between the localization and the judgment; in the other half localization and judgment agreed. Young considered that this experiment shows that "normal localization is possible on an auditory-motor basis despite the transposition of the ears" (**5**, 418). Young related his success in this test to timbre and intensity changes which resulted from changes in the position

of the ear trumpets of his pseudophone. Hard rubber trumpets facing slightly forward, as on Young's pseudophone, may be expected to provide much stronger directional cues than the normal pinnæ or the small ear-pieces used in the present experiment.

Normal localization by means of head movements, such as Young reported, was obtained by none of the observers in the present experiment. On the contrary, it was found that head movements, together with movements of the whole body, produced *more exact reversals*. This was repeatedly demonstrated by simple blindfold tests in a room with a radio at one side. If the observer was rotated a sufficient number of times to become disoriented and then was instructed to turn until he faced the radio, he would almost without exception face in the direction exactly opposite.

Young also reported a type of lateral reversal which was unaffected by movement. When he moved about in a room in the center of which a metronome was placed, the sound was localized on the opposite side and seemed to rotate as he walked. Under these conditions estimations of distance on the basis of intensity changes were incompatible with judgments of direction based on localization. Young states that in this situation "The task of approaching and 'finding' the phantom sound-object on the basis of auditory cues was impossible" (**5**, 403). In view of the similarity between the results obtained in this situation and those obtained when a radio was used in the present experiment, one is hardly justified in concluding that head and body movements ordinarily make possible correct localization when the pseudophone is being worn.

Purely auditory localization, or localization without visual knowledge of the position of the sound-producing object, was tested at the beginning and toward the end of the experiment. No change in localization of this type could be detected. Under conditions which did not permit the use of visual or other knowledge to assist in localizing the sound, lateral reversals were consistently reported. This held true not only for all reliable tests made in the laboratory but for all

other situations as well. Street sounds, conversation, and all the complex auditory stimuli experienced in the course of the experiment were reversed when the only localizing cues available were auditory. In view of the varied evidence on this point, in respect to which out findings were in complete accord with those of Young, it is clear that purely auditory localization is not easily changed as the result of transposition of the right and left auditory fields by means of the pseudophone.

The experiments of Stratton and Ewert on inversion of the visual field yielded a similar stability of purely auditory localizations. Stratton states that lateral changes in the localization of sound "came rarely except by strong effort of attention and voluntary vizualization" (**3**, 478). Ewert reports that "continuous inverted vision apparently has a neutral effect upon sound localization when the source of sound is in the . . . non-visual field. This indicates that during the period of retinal inversion, sound localization, in general, has no tendency to swing into line with visual localization" (**2**, 259–260). It appears to be definitely established that neither the pseudophone nor the inverting lenses used by Stratton and Ewert alter appreciably the cues used in purely auditory localization.

Localization of sounds within the field of vision, however, is much less resistant to change. When the observer can see the source of sound, it was found in the experiments of Stratton, Young, and Ewert, and in the present experiment, that the sound sometimes seemed to be coming from the position of the seen object. Young states that in his experiment this type of localization, which he terms 'visual-auditory localization,' *developed* during the period he was wearing the pseudophone to the point where it could not be distinguished from normal.[2]

[2] In a second article (**6**, 122), Young has stated that he no longer regards treatment of this problem in terms of *types of localization* as the best conceptual approach. He expresses the belief that instead it would be better to consider the *conditions of localization*, to be designated, in their various combinations, as *abc*, *acd*, *bcd*, etc. Until these conditions have been described in more specific terms, however, it seems necessary to continue to employ the concepts of the earlier article, with such supplementation as appears desirable in the light of the results of this experiment.

In the present investigation special attention was given to the possibility of verifying this 'development of visual-auditory localization' reported by Young. The critical question was formulated as follows: Is there progressive adaptation of auditory localization of this type? In contrast to the definite affirmative answer to this question made by Young, the results of this experimentation indicate that such normal localization as occurred did not *develop*. Instances of reversed localization which was unaffected by visual knowledge of the source of sound were reported at all stages of the experiment, and instances of normal localization were similarly distributed. Some of the most striking reversals occurred toward the end of the experiment.

The experimenters believe that any impression of progressive visual-auditory adaptation which is reported in an experiment of this sort needs to be checked very carefully. The number of reversals perceived appears to be dependent upon the amount of attention given to the localization of sounds. During the first days of wearing the pseudophone, the observer noticed unusual phenomena more than toward the end of the experiment. He soon learned that it was unsafe to rely on auditory localization for making motor adjustments of any sort without a deliberate correction of the immediate automatic response. After this was clearly realized his attention was directed less intensively to the matter of localization. The conditions of hearing then became comparable to those of ordinary normal hearing, when attention is given to the position of a sound only when one is suddenly startled or when one needs to use auditory localization to find an object. In such situations localization was invariably reversed, both at the beginning and at the end of the experiment. The localization of sounds in other situations was frequently neglected; this was done the more readily since auditory localization was no longer of functional significance. During the course of the experiment all the observers were vividly reminded that for the successful avoidance of traffic auditory localizing cues were of definitely negative value.

In the present experiment the normal localization which occurred was of such a special quality that it seems hardly

advisable to regard it as in any real sense a reorganization of visual-auditory perception. It may be better described as a temporary repression of the auditory cues. Visual knowledge of the position of an object, sometimes supplemented by long habituation to the coincidence in space of an object and its characteristic sound, may sufficiently over-ride the auditory cues so that the observer is not aware of any discrepancy between the visual localization of the object and the auditory localization of the sound. To conclude from this circumstance that auditory localization with vision has become normal seems definitely without warrant. A large number of observations support the view that the auditory components of 'visual-auditory localization' are altered to a very limited degree, if at all.

The importance of attention as a determinant of normal localization has already been noted. In many instances the observer might be unaware of a dissociation between visual and auditory 'thereness' so long as he was not thinking about localization, but when his attention was directed to the matter he would immediately realize that the sound was not coming from the place where the object producing it was located. Young reports a very similar experience; he states (**5**, 414) that "at the close of the experiment dissociation could be brought about through the control of observational attitude when the source was in the binaural axis." Young does not appear to have considered the possibility that the "development" of normal visual-auditory localization may represent a change in observational attitude rather than a fundamental reorganization of auditory perception.

A very important factor in the determination of normal localization is the degree of association between a sound and an object. *When a person talking to an observer laughed, or coughed, or changed his tone of voice from that which he customarily used, a clear reversal was immediately perceived, even though the localization just previously had seemed normal.* The strength of the association between the usual voice tones and the visual representation had at first overcome the auditory localizing cues. But a laugh or cough or the changed voice was not so definitely associated with an individual; and in the

localization of these the usual auditory cues played a dominant rôle, producing a reversal. Such adaptation as did take place in the localization of speech was thus limited to the normal tones of a familiar voice. The introduction of any change which weakened the association was followed by a clear reversal. Numerous other examples of the importance of associative factors were reported. Ordinarily sounds made by the observers themselves were normally localized. When the sound was made unintentionally, however, or when the action of the observer produced a sound at a distance, reversals were generally experienced. The accidental dropping of a wrench by one of the experimenters while working in the laboratory shop was heard on the side opposite that from which the sound was actually coming. On another occasion, when a piece of wood was thrown a short distance away, the noise which followed its striking the floor was unmistakably reversed, although the flight of the object was observed through its entire course.

It is probable that there are wide individual differences in the extent to which visual knowledge of the position of an object determines auditory localization. Young and Stratton appear to have experienced relatively few dissociations of auditory localizations from the visual representations which have long been associated with a particular sound. The visual-auditory localizations of Young which "developed so that the sight-sound coördinations were indistinguishable from those of normal life" depend, he believes "upon some visual cue which determines at the start the position of the source, and this cue may also determine the localization of the sound" (**5**, 427). Stratton writes (**3**, 478); "The sound of my footsteps conformed to the tactual and visual localization of my feet, and . . . in general the sight of the sound's place of origin carried with it the localization of the sound . . . ." The perceptions reported by two of Ewert's subjects, on the other hand, are like those obtained in the present experiment. In connection with testing sound localization in the visual field, Ewert (**2**, 266) found that "Inverted vision is a powerful distraction to sound localization. At first there is a tendency to localize the auditory-visual stimulus in terms of vision.

This tendency is entirely inhibited by 2 O's [of the three observing] in favor of localizing the stimulus in harmony with audition" despite the visual distraction. Wooster found similar differences in her subjects in respect to the dissociation of visual and auditory localization. These subjects looked through prisms which produced an angular deviation of the visual field of about 21 degrees. Under the conditions of her experiment, with the visual field shifted to the right, for most of the subjects the sound was at the beginning 'clearly assigned to a position definitely to the left of the object as seen' (4, 57). One subject, however, "reported that she was not at all conscious that the sound was to one side of the buzzer and did not even know until the very last the direction of the distortion" (4, 54). The report of a subject who is susceptible to visual suggestion to this degree will obviously differ from that of one who is able to maintain dissociation of visual and auditory localization. Possibly some of the difference between the amount of dissociation reported by Young and Stratton, and that reported by Ewert's subjects and in the present experiment may be attributed to this factor.

Further evidence of the influence of attention on localization was secured at a psychology seminar attended by the observers. It was found that if the speaker was followed closely localization was not consciously abnormal, even when the speaker was 45 or more degrees to one side of the median plane of the observer. When the observer gave primary attention to the localization of the speaker's voice, however, it was perceived as clearly reversed. If the observer was in a state of slight abstraction, the voices of the other members of the seminar were reversed whether the person speaking was visible to the observer or not. Observations of this character indicate that empathy plays an important rôle in the determination of localization. The set of the observer when following the speaker attentively appears to have been conducive to an identification of visual and auditory stimuli.

The experience of two observers at the cinema may be regarded as another example of the determination of localization by empathic response. At first a front-back reversal

was noted; the sound seemed to be coming from in back of them. After watching the screen a few minutes this reversal disappeared. Later when a crowd was collecting in the lobby behind them for the next performance, one of the observers localized the confused noise in the back of the theatre as coming from the screen.

The importance of empathy as a determinant of localization is further illustrated by a comparison between the radio and the human voice with respect to the frequency of occurrence of normal localizations. The radio was clearly reversed laterally with or without vision for all positions except when faced directly, and for this position a front-back reversal was frequently reported. The voice of a person with whom the observer was conversing, on the other hand, was usually localized normally when the person speaking could be seen clearly. When engaged in conversation it was necessary for the observer to consider very carefully from what direction the sound seemed to be coming to perceive reversed localization, and even with close attention a definite reversal was not always perceived. Normal localization of the voice was more likely to occur when the observer followed the movements of the lips and facial muscles. The larger number of reversals obtained with the radio may be attributed to the relatively slight opportunity it affords for empathic response, as well as to the fact that the association between an individual and his voice is at once stronger and more specific than the association between a radio and the varied auditory stimulation it is capable of producing.

Any attempt to explain normal localization under the conditions of this experiment by postulating the alternate operation of two exactly opposed systems of auditory localization—purely auditory and 'visual-auditory' appears to be unjustified. The term 'visual-auditory' seems unsatisfactory as a characterization of the normal localizations obtained in view of the large number of reversals reported throughout the present experiment when the source of sound was in the field of vision. The present authors are unable to subscribe to the notion of a gradual development of normal localization. Their position may perhaps be most easily made clear by

contrasting it with the interpretation which has frequently been attached to Young's experiment. Carr (1, 149), in his discussion of auditory space perception, states that Young "succeeded in reorganizing his auditory localizing habits so that they once more conformed to those of vision and contact. This auditory reorganization was effected largely through the influence of visual knowledge." In the present experiment nothing occurred which could be described as a *reorganization* of auditory localizing habits. Visual knowledge was ineffective in producing the impression of normal localization unless a strong association had been previously built up between a sound and an object. Even when this was the case normal localization was found to be unstable. A slight change in the stimulus, abstraction on the part of the observer, or greater attention to localization at the expense of empathy would cause the recurrence of clear reversals.

The normal localizations reported do not appear to be phenomena which are peculiar to the conditions of this experiment. On the contrary, it seems possible to relate them to the familiar auditory illusions associated with ventriloquism where localization appears to be a matter of empathic response based on long previous habituation to the association in space of given auditory and visual representations.

Both Stratton and Ewert seem to regard the instances of sight-sound coördination reported in their experiments as a temporary repression of auditory cues rather than as the development of a new type of localization. Stratton states (3, 478): "The cases in which the lateral localization of sound did accord with the new sight were no doubt due to a momentary strengthening of the influence of visual suggestion to such a degree that the usual dominant factor in lateral localization becomes subordinate." Ewert, in a test of sound localization in the visual field, actually found an increase in visual and auditory dissociation after the first day which permitted correct localization of a sound in spite of the inversion of the visual field. One of the conclusions he draws from his experiment applies equally well, *mutatis mutandis*, to the results of the present investigation. "The organism is able to adapt itself to a visual distraction, but this does not mean

that audition swings in line with vision nor that vision swings in line with audition. All this means is that the organism is able to inhibit various sensory distractions" (**2**, 259). In the present experiment it was found that under certain conditions the auditory localizing cues were 'inhibited' to such a degree that the subject was aware of no discrepancy between visual and auditory localization.

Summary.—In this experiment two observers wore pseudophones continuously, except at night, for one week; a third observer wore a pseudophone for three days. Purely auditory localization (without the aid of visual cues) remained reversed throughout the experiment. No progressive adaptation of so-called 'visual-auditory localization' was found. Movements of the head and body resulted in more exact reversals, rather than in normal, or objectively correct, localizations. The authors believe those instances of normal localization which occurred during the experiment are to be attributed to a temporary repression of auditory localizing cues resulting from the counter-influence of association and empathy, rather than to visual knowledge of the position of the object. Reversal with vision was invariable except when association and empathy appeared to be involved.

The results of the present experiment are definitely opposed to the view that auditory localization can be radically altered within the space of a few days. Such normal localization as occurred under the conditions of the experiment does not seem to represent a fundamental reorganization of auditory habits.

(Manuscript received June 2, 1936)

REFERENCES

1. CARR, H. A., *An Introduction to Space Perception*, New York, Longman's, 1935.
2. EWERT, P. H., A study of the effect of inverted retinal stimulation upon spatially coördinated behaviour, *Genet. Psychol. Monog.*, 1930, **7**, nos. 3 and 4.
3. STRATTON, G. M., Vision without inversion of the retinal image, *Psychol. Rev.*, 1897, **4**, 341–360 and 463–481.
4. WOOSTER, M., Certain factors in the development of a new spatial coördination, *Psychol. Monog.*, 1923, **32**, no. 4.
5. YOUNG, P. T., Auditory localization with acoustical transposition of the ears, *J. Exper. Psychol.*, 1928, **11**, 399–429.
6. YOUNG, P. T., The role of head movements in auditory localization, *J. Exper. Psychol.*, 1931, **14**, 95–124.

STUDIES IN RETROACTIVE INHIBITION: IX. RETROACTIVE INHIBITION, REPRODUCTIVE INHIBITION AND REMINISCENCE

BY JOHN A. McGEOCH
Wesleyan University

FRED McKINNEY AND HENRY N. PETERS
University of Missouri

Problem

The phenomenon of retroactive inhibition has traditionally meant the decrement in retention of an original learned activity following interpolated learning. The phenomenon of reproductive inhibition has meant the decrement in the retention of an original learned stimulus-response sequence when one term of the original pair has, between learning and the testing of retention, been connected with another and different term.

The 2 phenomena have been regarded as different from each other in at least 2 ways, both of which go back to differences in the experimental operations by means of which they are measured. (1) Retroactive inhibition has been thought of in terms of series of items or of stimulus-response connections; in virtually all of the experiments both original and interpolated activities have been in serial form. Reproductive inhibition has been thought of in terms of single pairs of items or of stimulus-response connections; in the experiments the units of material have been pairs, with the interpolated or interfering pair having one term identical with and one term different from the corresponding term of the original pair. (2) The phenomena have also been regarded as disparate because of the clear experimental emphasis on the common term in the reproductive inhibition work and the absence of such emphasis in the retroactive inhibition work. The reproductive inhibition paradigm squarely affirms this identity of one term; the retroactive inhibition paradigm contains no necessary reference to any form of identity.

The demonstration, however, of the basic importance of similarity between original and interpolated lists as a determining condition of retroactive inhibition suggests that retroactive and reproductive inhibition may be closely related or even identical phenomena. The common term required by the reproductive inhibition paradigm may be regarded as a limiting case of both the formal and meaningful dimensions of similarity.

If the two are essentially the same, and if connection of original and interpolated activities with a common term or with common terms is a major factor, experimental conditions which differ only with respect to the clear presence of a common term should show the greater inhibition for the condition containing this term. The amount of the difference should, moreover, aid interpretation of the similarity problem. The experimental data on similarity in retroactive inhibition permit inferences about the outcome, but direct experimental measurement is to be preferred to inference.

EXPERIMENT I

Procedure.—We have made by the method of group measurement certain preliminary comparisons between the recalls of paired associates under conditions designed to yield retroactive inhibition and under those designed to yield reproductive inhibition. These comparisons have yielded data, also, upon reminiscence and, indirectly, upon associative inhibition.

The learning materials were lists of 20 pairs of Chinese-English words.[1] Each list used in the experiment was studied by the subject for 4 minutes. The amount learned during this period was tested immediately by the method of right associates in which the Chinese words were presented in an order different from that of the learning and the subjects were instructed to write the English equivalents. They were told that they would have 4 minutes in which to do this. The procedure, thus far, was followed with every list. The other conditions and the composition of the lists will now be described individually.

Under Condition 1 (*rest*) the test was repeated after an interval of 10 minutes from the completion of the first test. At the second or delayed recall test the Chinese words were in the same order as at the first test and the subjects were again given 4 minutes in which to write the English equivalents. The ten-minute interval was filled, save for the 30 seconds, at the end of the interval, devoted to the instructions for the second recall, with the non-learning activity of cancelling numbers (5's) on a large mimeographed number sheet. This activity was designed to prevent rehearsal. Hereafter the immediate recall, which is a test of learning, will be designated as *learning* and the delayed recall as *recall*.

[1] The words used are from *Notes for Instructors* to accompany W. S. Foster and M. A. Tinker, *Experiments in Psychology*. N. Y.: Holt, 1929, pp. 20–21.

Under Condition 2 (*retroactive inhibition*) the subjects turned at once, after the completion of the first test of the first list, to the study of a second list composed of 20 different pairs of Chinese-English words. The study and test of this interpolated list occupied 8 minutes. The remaining 2 minutes of the ten-minute interval corresponding to that of the rest condition was filled with cancellation of numbers as under rest (1 minute, 30 seconds), and by the instructions for the second recall (30 seconds). In Condition 3 (*associative and reproductive inhibition*) the procedure duplicated that of Condition 2. The interpolated list, however, contained the same Chinese words as had the original list and in the same order but with different English equivalents. The Chinese words were presented at the test in the same changed order as the one used for the original list. The English words in the original list were verbs; those in the interpolated list were nouns. It was necessary to inform the subjects of this in each case since at the delayed recall of the original list some exact designation of the words to be recalled, over and above the statement that they were the words in the first list, was desired. Condition 3 differs formally from Condition 2 only in the fact that the Chinese words were the same in both lists and that, therefore, the pattern of conditions supposed to be clearly favorable to the appearance of associative and reproductive inhibition is provided.

In Condition 4 the list which served as the interpolated list in Condition 2 was studied and tested once and in Condition 5 the same was done with the interpolated list of Condition 3. The results yield a measure of the influence exerted by the original lists upon the interpolated ones, and provide a measure, by a comparison of the learning of A–C and D–E alone and following A–B, of the amount of associative inhibition.

The conditions are summarized in the following schematic outline in which A–B always stands for the same list of 20 pairs, D–E stands for a second list with no items in common, and A–C represents a second list in which the Chinese words are the same as those in A–B.

Condition	Original List	Interpolation	Delayed Recall
1	A–B	Rest	A–B
2	A–B	D–E	A–B
3	A–B	A–C	A–B
4	D–E	——	——
5	A–C	——	——

A different group of University of Missouri students[2] was employed under each condition. Each condition was gone through by the entire group at the same time. The lists of pairs and the lists of Chinese words for the tests were mimeographed, each on a separate sheet. At the beginning of the experiment, each subject was given all of the materials required by the experiment. The sheets were clipped together in order and at each successive step in the experiment the subject had only to place the sheet at which he had been working face down at the bottom of his packet of materials, turn over the next sheet at the signal and proceed. It was possible, thus, to follow rigidly the time schedule of the experiment. Precautions were taken against copying and communication from subject to subject and it is believed that errors from this source are negligible. The groups under Conditions 1 and 2 were in charge of one

[2] All subjects were from sections of General Psychology.

experimenter; those under Conditions 3, 4 and 5 were in charge of a second experimenter. The two experimenters had very carefully gone over the procedure and had agreed upon each detail.

Results.—The mean number of right associates and the sigma of the mean of each list and under each condition are given, together with the percentage of the original lists recalled after the interval in Conditions 1 to 3, in Table I.

TABLE I

MEAN RECALL SCORES, SIGMAS OF THE MEANS AND PERCENTAGES OF DELAYED RECALL

Condition	N	Original List Learning	Interpolated List Learning	Original List Recall	Percent Recall Orig. List
1. (A–B—A–B)	56	10.71±.54		11.48±.54	107.2
2. (A–B; D–E; A–B)	75	10.52±.50	13.62±.51	9.81±.56	93.3
3. (A–B; A–C; A–B)	51	10.76±.59	14.55±.61	8.11±.67	75.4
4. (D–E)	47	12.34±.59			
5. (A–C)	50	14.30±.66			

The original list (A–B) under Conditions 1 to 3 is learned to an almost identical degree by the 3 different groups, when learning is measured by the number of right associates at the first test. The interpolated lists (D–E and A–C) have higher means than do the original lists. We shall, however, disregard this for the moment and proceed to a comparison of the delayed recalls of the original list. Although the means of the 3 groups which learned this list are nearly equal at learning, their corresponding means at recall show considerable differences. The difference of 1.67 between the means at recall under Conditions 1 and 2 has a sigma of ± .78; that of 3.37 under Conditions 1 and 3 has a sigma of ± .86; and that of 1.70 under Conditions 2 and 3 has a sigma of ± .87. Only the second of these differences is 3 times its sigma, but the probability of a true difference in the 2 other cases, where the critical ratio is roughly 2, is relatively high.

The delayed recall under the rest condition is slightly higher than the immediate recall or learning; the percentage of delayed recall is 107.2. When an interpolated list composed of entirely different items is studied and recalled be-

tween the learning and delayed recall of the original list, the percentage of recall of the original list is decreased to 93.3. The interpolation of a list composed of the first members of the pairs in the original list and of new second members decreases the percentage of recall of the original list still further, *i.e.* to 75.4. If we compute from these percentages the percentages of decrement, they are found to be 13.0 under the retroactive inhibition condition and 29.7 under the reproductive inhibition condition.

Lists D–E and A–C, when learned alone, yield for some reason, which need not concern us here, larger means than does list A–B. List D–E learned alone (Condition 4) has a mean which is 1.28 ± .78 lower than the same list learned after list A–B, which implies a practice effect from A–B to D–E. The difference, however, between list A–C learned alone and the same list learned after A–B is very small (.25 ± .90). This virtual equality under conditions which formally permit a practice effect such as occurred from A–B to D–E may be ascribed to the presence of the common element A in the lists A–B and A–C. This difference in the amount of practice effect is, thus, evidence for a small amount of associative inhibition. The presence of the common item in each pair has, however, done no more than to reduce the influence of practice from the first list to the second below the amount which the practice effect from A–B to D–E would lead us to expect. The latter would lead us to expect a practice effect of 1.28 and the obtained practice effect is 1.03 lower than this. These differences have a low statistical reliability, but even if we disregard the fact that they are unreliable, the difference of 1.03, plus an unknown but certainly small correction which would need to be made in consideration of the higher mean of the A–C list when learned alone, represents not far from 10 per cent of associative inhibition.

The presence of the common item has had, in spite of its small influence upon the learning of A–C, a more important influence upon the recall of the A–B items. That is, although the amount of associative inhibition, or inhibition in the

direction A–B to A–C, is small, the amount of reproductive inhibition, or inhibition in the direction A–C to recall of A–B, is of considerable size. The appearance of reproductive inhibition is, under these conditions, therefore, not dependent upon an equal amount of associative inhibition.

We have already mentioned the fact that under the rest condition recall after the interval is slightly larger than it is immediately after study. The percentage of recall is 107.2. This gross recall of more than 100 percent is another illustration of the failure of a single recall to provide a completely stable measure and is conventionally called reminiscence.[3] Under Conditions 2 and 3 gross reminiscence disappears. The 2 interpolations have eliminated it in so far as the means can reveal it. Mean recall score is not, however, an adequate indication of the presence or absence of reminiscence[4] and we have analyzed the data in terms of the number of subjects showing reminiscence. The results of this analysis (Table II) bear upon the influence of the interpolated conditions as well as upon reminiscence.

There are differences between the mean learning scores of the subjects who show the 3 classes of change from learning to recall, but in the absence of any consistency in the direction of the differences it cannot be concluded that any particular kind of change is a function of a difference in degree of learning as measured.

The meaning of the data of Table II may be summarized in a few statements. (1) Reminiscence, measured in terms of the number of subjects having a larger score at recall than at learning, appears under each of the 3 experimental conditions. (2) The percentage of subjects showing reminiscence decreases from Condition 1 to Condition 3. The 2 conditions of interpolation decrease percentage of reminiscence below that found under rest, but the decrement produced by the reproductive inhibition condition is much the larger. (3) The

[3] *Cf.* W. Brown, To what extent is memory measured by a single recall? This JOURNAL, 1923, **6**, 377–382. It should be understood that we mean by 'reminiscence' only an increased score from Recall 1 to Recall 2 when the 2 recalls are made by a single subject.

[4] *Cf.* G. O. McGeoch, The conditions of reminiscence. *Amer. J. Psychol.*, 1935, **47**, 65–89.

TABLE II

LEARNING AND RECALL SCORES OF SUBJECTS WHO GAIN, LOSE OR ARE CONSTANT FROM LEARNING TO RECALL

Change from Learning to Recall	No. of Ss	Percent of Ss	Original Learning	Interpolated Learning	Recall Original	Percent of Recall
Condition 1 (Rest)						
Gain.....	29	51.8	9.89	——	12.55	126.9
Loss.....	17	30.3	11.23	——	9.23	82.2
Zero.....	10	17.9	12.20	——	12.20	100.0
Condition 2 (Retroactive Inhibition)						
Gain.....	26	34.7	11.42	15.19	13.38	117.2
Loss.....	38	50.7	10.58	12.66	7.84	74.1
Zero.....	11	14.6	8.18	13.27	8.18	100.0
Condition 3 (Reproductive Inhibition)						
Gain.....	10	19.6	10.50	14.60	12.90	122.9
Loss.....	41	80.4	10.83	14.53	6.95	64.2
Zero.....	0	——	——	——	——	——

reminiscence, as measured by percentage of recall, yielded by the subjects who show reminiscence at all, is slightly greater under Condition 1 than under Condition 3 and is somewhat smaller under Condition 2 than under the others. The conditions of interpolation have not greatly affected the amount of reminiscence shown by those who show it at all. (4) The percentage of subjects showing a loss after the interval increases from Conditions 1 to 3 until the percentage is high under Condition 3. The percentage showing zero change is relatively low under the 2 first conditions and is zero under the third. (5) The percentage of items recalled by the subjects who show a loss also declines from Condition 1 to Condition 3.

When the influence of the interpolated lists is computed only in terms of the mean recall score of the group, certain important aspects are, thus, obscured. A more detailed analysis reveals that the decrement brought about by the reproductive inhibition condition is larger than that produced by the retroactive condition not only in terms of total score,

but also in terms of the number of subjects affected and in terms of the extent to which these subjects are affected.

Experiment II

We have reproduced in a second experiment Conditions 1 and 3 of Experiment I. We have reproduced, also, Condition 2 with the single exception that List D–E in the scheme of that experiment has been used as the original list and another list, A–F, as the interpolated list. In List D–E the English words are nouns and in List A–F they are verbs. This Condition 2 has been repeated with another group using Lists D–E and A–C, in both of which the second terms (English words) are nouns. We shall call this Condition 2A.

The data, corresponding to those of Table I, are given in Table III. The mean learning scores are sufficiently close

TABLE III

Mean Recall Scores, Sigmas of the Means and Percentages of Delayed Recall

Condition	N	Original List Learning	Interpolated List Learning	Original List Recall	Percent Recall Orig. List
1. (A–B—A–B)	81	10.83±.43	——	10.60±.49	97.9
2. (D–E; A–F; D–E)	118	11.68±.42	11.07±.46	10.86±.44	93.0
2A. (D–E; A–C; D–E)	104	12.14±.46	14.31±.43	10.89±.50	89.7
3. (A–B; A–C; A–B)	85	8.84±.44	13.08±.47	6.38±.43	72.2

to those of Experiment I to guarantee the adequacy of the repetition of the conditions in every case save Condition 3. Here the means at both original and interpolated learning are lower in Experiment II than in Experiment I. The academic grades of the students working under Condition 3 are somewhat lower than the grades of the students who served under the other conditions. It may be that the hour at which this section met, or some other selective factor, has brought it about that this group is of lower general ability than the others. This may account for the lower mean learning scores. Since, however, we are interested primarily in percentages of recall rather than in absolute scores, and since these percentages agree with those of Experiment I, these

lower means under Condition 3 do not seriously affect interpretation.

The percentages of recall under the 3 main conditions stand in the same relations as did the corresponding percentages of Experiment I. The rest condition of Experiment II shows no gross reminiscence but recall is close to 100 percent. Retroactive inhibition 2 has almost exactly the same percentage of recall as in the first experiment, retroactive inhibition condition 2A has only slightly less, and the reproductive inhibition condition differs only slightly in percentage of recall from the first experiment. Since the mean at original learning in Condition 3 is less than the means under the other conditions, and since this is apparently a less able group, the lower mean for the interpolated list (A–C) than for this list under Condition 2A cannot be construed to mean the presence of associative inhibition. On the contrary, the mean of the interpolated list is so much higher than that of the original list that positive rather than negative transfer is indicated. It is probable that we have here a considerable amount of reproductive inhibition without the appearance of any considerable amount of associative inhibition.

No gross reminiscence has appeared in Condition 1, but in view of the emergence of relatively large amounts of reminiscence when the data of the individual subjects are examined, it is worth while to study these data in the same way. The results for corresponding conditions (Table IV) parallel very closely those of Experiment I. Some subjects show reminiscence under each condition, and the percentages showing reminiscence decline from Condition 1 to Condition 3. The percentages showing it are smaller than in Experiment I, but the percentages of recall by those who do show reminiscence are nearly the same under the 3 conditions and do not differ importantly from the corresponding percentages of the first experiment. The percentages of subjects showing neither gain nor loss are larger in this experiment, but decline from Conditions 1 to 3. The percentages of subjects showing a loss are also larger and they increase from Conditions 1 to 3. The percentages of recall by subjects

TABLE IV

LEARNING AND RECALL SCORES OF SUBJECTS WHO GAIN, LOSE OR ARE CONSTANT FROM LEARNING TO RECALL

Change from Learning to Recall	No. of Ss	Percent of Ss	Original Learning	Interpolated Learning	Recall Original	Percent of Recall
Condition 1 (Rest)						
Gain.....	22	27.2	11.18	——	13.64	122.0
Loss.....	36	44.4	10.53	——	8.50	80.7
Zero.....	23	28.4	11.00	——	11.00	100.0
Condition 2 (Retroactive Inhibition)						
Gain.....	24	20.3	10.58	10.38	13.13	124.1
Loss.....	69	58.5	11.69	11.06	9.41	80.5
Zero.....	25	21.2	12.68	11.80	12.68	100.0
Condition 3 (Reproductive Inhibition)						
Gain.....	8	9.4	8.25	11.50	10.00	121.2
Loss.....	73	85.9	9.12	13.49	6.07	66.6
Zero.....	4	4.7	5.00	8.75	5.00	100.0

showing a loss are the same under Conditions 1 and 2 and are considerably higher than the corresponding percentage under Condition 3.

DISCUSSION AND CONCLUSIONS

We may now summarize the results of the 2 experiments and discuss their implications. (1) The learning of an interpolated list of Chinese-English words produces a small decrement in the retention after 10 minutes of an original list of Chinese-English words, when the 2 lists have no common terms. (2) When, however, each pair of words in the interpolated list contains the same Chinese word which appeared in the pair at the corresponding position in the original list, the decrement in recall is considerably greater. These 2 statements hold both for unanalyzed mean scores and percentages of recall and for more completely analyzed scores.

When the data are analyzed in more detail, it is found that the influence of both forms of the interpolated learning

has been (*A*) to reduce the percentage of subjects showing reminiscence, (*B*) to decrease the percentage of subjects showing no change in score from learning to recall, and (*C*) to increase the percentage of loss on the part of those subjects who show a loss.[5] These effects are uniformly larger under Condition 3, where the pairs in the 2 lists have common terms than in Condition 2, where the pairs have no common terms. The results from this more detailed analysis of the data specify importantly and strengthen greatly the general conclusions of the paper. The conclusions hold not only for total scores but also for number of subjects affected by the interpolations and for the extent to which they are affected.

(3) The reproductive inhibition measured under Condition 3 appears without an equal amount of associative inhibition. That is, the decremental influence of the learning of A–C upon the recall of A–B is considerable, although there is either a small decrement in the learning of A–C from the previously learned A–B or none at all. This demonstration of the appearance of reproductive inhibition without a correspondingly large amount of associative inhibition is a significant fact in the field of inhibition and has important implications for the relations between learning and retention. It is also in harmony with the fact that overt transfer in retroactive inhibition is much greater from interpolated list to delayed recall of original list than from original to immediate recall of interpolated.[6] It needs, however, to be checked under conditions which permit a more precise measure of practice effect and hence a more precise statement of amount of associative inhibition.

(4) A considerable percentage of the subjects in the rest condition of both experiments show the phenomenon of reminiscence. It appears with at least a few subjects in

[5] An exception occurs in Condition 2, Experiment II, where the greatest effect is in the increased percentage of subjects losing and the decreased percentage of those gaining or remaining constant. Condition 2 shows no increase over Condition 1 in the percentage of loss by subjects showing a loss. In view of the effects first mentioned, however, this exception is relatively unimportant.

[6] *Cf.* F. McKinney and J. A. McGeoch, The character and extent of transfer in retroactive inhibition: Disparate serial lists. *Amer. J. Psychol.*, 1935, **47**, 409–423.

every condition. The appearance of reminiscence in college students and with disparate materials supports the conclusion that the phenomenon is not specific to subjects below early adulthood and to meaningful materials. The decremental influence of interpolated learning upon the number of subjects showing reminiscence is also interesting. The fact that the percentage of recall by subjects who show reminiscence approaches a constant is important for a theoretical interpretation of the phenomenon of reminiscence.

The results clearly show that, under these conditions, an interpolated list of paired associates, each pair of which contains the same initial term as the corresponding pair in the original list, produces more inhibition than does an interpolated list of similar kind but without the common term.[7] There seems little reason, in the light of these and the other data on retroactive and reproductive inhibition, to regard the two as disparate phenomena. The common term in the paired associates of this experiment may be considered as belonging to one or more dimensions of similarity and to represent an obvious link between reproductive inhibition and the experiments upon retroactive inhibition as a function of similarity. The hypothesis that the 2 forms of inhibition are fundamentally identical and that the common term is classifiable in the category of similarity would lead us to

[7] In connection with this result it is necessary to consider the question whether in the reproductive inhibition condition the lessened recall may have been a function of inability on the part of the subjects to identify the units called for in Recall 2. The Chinese words were the same in both original and interpolated lists, and it might be that at the second recall of the original list the subjects were uncertain which English words were expected. We believe this to be possible but very improbable. (1) The instructions were clear and no subject expressed overtly any uncertainty about the list meant at Recall 2. (2) If inadequate identifiability has been an important factor, more words from the interpolated list should appear in Recall 2 of the original list under the reproductive inhibition condition than under the retroactive inhibition condition. We have analyzed the recalls in Experiment I. There is an average of .04 more confusions of the words in the 2 lists under the reproductive inhibition than under the retroactive inhibition condition. That is, an average of .04 more words from the interpolated list appears in the second recall of the original list under one condition than the other. This insignificant difference gives no support to an interpretation of the results as a function of inadequate identifiability.

expect the obtained result that Condition 3 produces more inhibition than Condition 2.[8]

The chief question concerning this interpretation arises in the cases where the original and interpolated lists have no obvious common term and no measured similarity of the kind used in the experiments on the similarity problem. Are there in these cases the common terms required by the reproductive inhibition paradigm? It is a reasonable hypothesis that there are. Other studies of this series have suggested what some of these common terms may be and it is entirely possible that such terms may be found wherever so-called retroactive inhibition appears.

(Manuscript received February 15, 1936)

[8] It is a pleasure to acknowledge our indebtedness for many critical comments to Dr. Arthur W. Melton of the University of Missouri.

THE ALLEGED RETROACTIVE EFFECT OF VISUAL STIMULI SUBSEQUENT TO A GIVEN RESPONSE

BY FREDERICK A. COURTS

Reed College

As a disproof of the law of effect, Tolman, Hall and Bretnall (9) have offered evidence that "in a trial and error situation" . . . "an accent or emphasis (such as bell or shock) upon the correct responses will tend to favor learning, whereas an accent or emphasis upon the wrong responses will tend to hinder learning." This principle (known as the Law of Emphasis), together with a Law of Motivation and a Law of Disruption, is offered as a substitute for Thorndike's Law of Effect.

The "Law of Emphasis," however, is not necessitated by the results of their experiment. On the contrary, their data may be interpreted as additional proof of the law of effect. In telling the subject, after he had run the 'punchboard maze' three times, inserting the stylus in every hole, that he was to select those holes which resulted in the ringing of the bell (or those which resulted in the bell and the shock), these two factors (the bell and the shock) were set up as goals. They became indicators of the correctness of response. Thus, the bell and the shock may have induced in the subject a 'satisfying' after-effect, such as that which accompanies success. The fact that performance was better when the bell and the shock accompanied the correct responses than it was when they occurred with the incorrect ones may thus be taken as a confirmation of the hypothesis that a 'satisfying' after-effect strengthens a connection more than an 'annoying' after-effect weakens that connection. The experiment of Tolman, Hall and Bretnall appears to be inconclusive.

The results of this experiment have been criticized adversely by Goodenough (1) on the grounds that the obtained

differences in performance may have been due to differences in the learning ability of the groups rather than to the influence of the so-called 'emphasizers.'

The experiment was repeated by Muenzinger (**6**), who concludes that the differences between performances under the various conditions are of slight statistical significance.

Furthermore, Muenzinger (**5**) has found that with rats . . . "moderate shock in the visual discrimination habit has a weakening or inhibiting effect if it accompanies the wrong response; it has a strengthening or facilitating effect if it accompanies the right response." His results, moreover, lend support to our contention that the experiment of Tolman, Hall and Bretnall is merely an additional demonstration of the fact, recognized some time ago by Thorndike (**8**), that rewards are much more effective in strengthening a connection than punishments are in weakening that connection.

On the other hand, Hulin and Katz (**4**) have recently repeated and confirmed the 'punchboard maze' experiment. However, since their experimental procedure was an exact duplication of the original investigation, their results are no more conclusive than those of Tolman, et al.

Stephens (**7**), in an ingenious experiment, has adduced evidence which may be interpreted as a confirmation of the 'Law of Emphasis.' In this experiment the subject attempted to hit, with a stylus, a particular spot on a horizontal plate which he could not see. Some responses were followed by the flashing of a lighted R, others by a W. These signals meant right and wrong response respectively. Other responses were followed by the flashing of both letters together. In terms of the distance by which a hit deviated from the point struck in the preceding trial, the responses followed by a signal of R had a greater tendency to be repeated than those followed by the flashing of both letters. These (the ones followed by both) were, in turn, more likely to be repeated than were the responses followed by a signal of W. All of the responses followed by either or both of the letters were found to be more apt to be repeated in the subsequent trial than were those which were followed by no signal. Stephens suggests that

"the mere fact that 'something' happens after a response tends to strengthen the underlying connection, and that a rewarding something may add to the strengthening influence and a punishing influence detract from it." Thus, he explains the difference in effectiveness of 'satisfying' and 'annoying' after-effects.

Stephens' subjects were instructed to inform the experimenter each time that the light flashed. They were to say 'right' or 'wrong' accordingly as the R or W was illuminated. When both letters appeared, the subjects were to say 'both.' *There was no attempt to determine what meaning the subjects attached to the flashing of both letters.* It is possible that the emphasizing factor, in this experiment, as well as in that of Tolman, Hall and Bretnall, played the dual role of 'emphasizer' and informative cue. If the appearance of both letters informed the subject, with even a slight degree of certainty that his response was right, it must be concluded that Stephens' interpretation of this experiment is not justified. The fact that responses followed by the flashing of a light which (from the experimenter's point-of-view) had no meaning, tended to be repeated more often that those responses which were not followed by the flashing of the light may be taken as evidence that the subjects tended to view the light as a signal of correct response. Hall (**3**) has indicated that in the absence of definite information of success or failure there is a tendency for human subjects to evaluate their responses as being right or wrong. Stephens' signal of both lights may have meant 'almost right' or 'maybe right' to the subject.

Inasmuch as the experimental procedure in Stephens' investigation, as well as in that of Tolman, Hall and Bretnall, has not provided an adequate control for this factor of informative cue, it is necessary to invoke nothing more than Thorndike's law of effect as an explanation of the results of these experiments. Thus Tolman's Law of Emphasis and Stephens' statement that "the mere fact that 'something' happens after a response tends to strengthen the underlying connection" appear to be superfluous.

The present experiment represents an attempt to avoid this difficulty by making it impossible for the 'something' which follows the response to acquire the character of an informative cue. Only in this way is it possible to determine the effect of an informationless 'something happening' after a response.

Apparatus.—A device similar to Jastrow's memory apparatus was used. It was arranged so that visual material could be presented simultaneously through two apertures one-half inch apart. Below these apertures were two keys, each of which was connected in series with an electric light of low candle-power which was visible only to the experimenter. Above the apertures there was a stimulus light, controlled by a relay which could be switched into circuit with either of the subject's keys by means of a switch operated by the experimenter. The stimulus light consisted of a three candle-power lamp, in circuit with 6 volts a.c. For each flash of this light, the circuit was closed for approximately 35σ. The rate of exposure of the material was controlled by an electrically operated shutter placed in front of the aperture. Each exposure was of six seconds duration. There was an interval of four seconds between exposures.

Standardization of Material.—Glaze (1) has compiled a list of one-hundred nonsense syllables which have 'no associative value.' In order to secure material of uniform difficulty we have selected sixty syllables from this list in the following manner. The one-hundred syllables were presented one at a time to fifteen subjects. The number of subjects responding with an associated word within fifteen seconds after presentation was noted for each syllable. The mean number of associations was found to be 6.4 Sixty syllables were selected from those which had a value of 7 associations or less. These were paired at random.

In order to avoid difficulties resulting from having the illumination follow consistently, for any one subject, the correct or the incorrect responses, ten of these pairs were designated as light-right, ten as light-wrong, and ten as no-light syllables. With five pairs of each group the correct choice was on the right-hand side. This classification was determined by chance.

The mean associative value of each of the groups is shown in Table I.

TABLE I

MEAN NUMBER OF ASSOCIATIONS FOR EACH GROUP

	Mean	σ dist.	σ m.
Light-Right.....	4.40	1.685	.377
Light-Wrong.....	4.55	1.284	.287
No-Light.....	4.305	1.112	.249

That these groups are comparable in terms of the number of associations given for each syllable is shown by the following values:

TABLE II

SIGNIFICANCE OF THE DIFFERENCE BETWEEN THE MEAN ASSOCIATIVE VALUES OF THE THREE GROUPS OF SYLLABLES

Groups	Diff.	σ diff.	$\left(\frac{\text{C.R. diff.}}{\text{diff.}}\right)$	P. Chances in 100 of this being a true diff.
L–R minus No–L.	.095	.450	.21	58
L–W minus L–R.	.150	.474	.32	62
L–W minus No–L.	.245	.377	.65	74

In order to prevent the subjects from learning the task as a series of right and left choices the pairs of syllables were presented in a different order for each trial. Each of the pairs of syllables was given a number. Twenty-four chance distributions of the thirty numbers were obtained. Each presentation of the series followed, for each subject, a different one of these chance orders.

Procedure.—The following written instructions were given each subject immediately before the experiment:

> This is a test of your learning ability. A series of pairs of nonsense syllables will appear, one pair at a time, in the small aperture which is directly in front of you. Each pair will be exposed for six seconds. There will be four seconds between exposures.
>
> The key on your left corresponds to the syllable on that side. The right-hand key corresponds to the syllable on the right.
>
> As soon as you see the syllables, press the key on the left. You will then be told if the syllable on that side is right or wrong. After hearing the word 'right' or 'wrong,' press the key on your right. You will then be told if the syllable on the right side is right or wrong. Then, wait for the next pair to appear.
>
> There is nothing about the syllables to indicate whether they are right or wrong. *You must try to remember the correct choice for each pair.*
>
> After the entire list has been presented, you will have a short time to rest.
>
> The list will then be repeated. The pairs will be the same as before, but the order will be changed.
>
> In this trial you are to press the key for the syllable which was correct in the preceding trial.
>
> *This time you will not be told whether you are right or wrong, but a record of your mistakes will be kept.*
>
> The practice and the test periods will be alternated in this way until you have learned the correct choice for each pair. You must go through the entire series once without a mistake.

The subject was instructed to execute both the correct and the incorrect response for each pair of syllables during the practice trial. For ten pairs of syllables the light flashed after the correct response during this trial. In the case of ten others the light followed the incorrect response. No illumination followed the responses to the remaining ten pairs of syllables. After each response, during the practice trial, the experimenter informed the subject of the correctness of his choice by saying 'right' or 'wrong.'

The subject attempted to repeat, in the test period, those responses which had been correct in the practice trial. During the test period there was no illumination, and the subject was given no information concerning the correctness of his responses.

Since there is only one light, since it flashes an equal number of times after the right and the wrong response, and since these conditions occur in random order, we have secured an *informationless* 'something happening' after a response.

In the 'punchboard maze' experiment it was necessary for the subject to execute all of the correct responses for each trial. Since this was not necessary for the incorrect responses, no one of these responses occurred for every trial. All of the correct responses did, however, occur for each trial. Thus the total frequency of occurrence of the correct responses (and the 'emphasizer' which accompanied them) was greater than that of the incorrect responses. If 'emphasis' does influence learning, the superiority of performance when the bell and the shock followed the correct responses may have been due to the greater frequency of this condition. In the present experiment this difficulty was avoided by alternating test and practice periods, where the subject executed all of the correct and the incorrect responses, and the conditions of light-right, light-wrong, and no-light occurred an equal number of times.

In view of the method of selecting the nonsense syllables, and the chance arrangement of these syllables, we are justified in assuming that the light-right, the light-wrong, and the no-light groups of syllables were of equal difficulty for the subjects.

We should find, then, according to the Law of Emphasis and Stephens' hypothesis a larger number of errors for the light-wrong group of syllables than for either of the other two groups. There should also be fewer errors for the light-right group than for the no-light group.

Results.—The average number of errors per trial for each group of syllables is shown in Table III.

TABLE III

AVERAGE ERRORS FOR THE THREE GROUPS
N=36

Trial	Light-Right	Light-Wrong	No-Light
1	4.00	4.19	4.64
2	3.06	3.22	3.28
3	2.75	2.64	3.19
4	2.19	2.03	2.06
5	1.11	1.72	1.81
6	1.03	1.25	.94
7	1.03	.89	.92
8	.67	.69	.69
9	.31	.36	.33
10	.19	.22	.11
11	.00	.06	.00
12	.00	.00	.00

The values presented in Tables IV and V indicate that the small differences between the number of errors under the three conditions are of low reliability. Only in the first

three trials are there differences large enough to indicate the presence of any but chance factors. In these trials more errors occurred on the no-light syllables than on those of either the light-right or the light-wrong group. In these first three trials, illumination seems to aid learning when it follows either the correct or the incorrect responses, but only in a degree slightly above that of a chance expectation. If, as these data may be interpreted, the illumination does affect performance during the first stages of learning, the effect is so slight that other, more important factors, obscure its influence as learning progresses. Moreover, the slight differences which do occur during these trials are in a direction which would lend little comfort to either Tolman's Law of Emphasis or Stephens' hypothesis.

TABLE IV

MEANS AND S.D.'S OF THE ERRORS FOR EACH GROUP

Trials	M.	σ dist.	σ m.
	Light-Right		
1– 3	9.53	3.44	.574
4– 6	5.19	3.55	.591
7–11	2.19	2.06	.343
1–11	16.61	8.99	1.489
	Light-Wrong		
1– 3	10.06	2.88	.479
4– 6	5.00	3.35	.591
7–11	2.22	3.73	.622
1–11	17.33	8.67	1.445
	No-Light		
1– 3	11.11	3.41	.588
4– 6	4.89	4.26	.710
7–11	2.06	3.34	.557
1–11	18.00	9.30	1.550

The high reliabilities shown in Table VI show that whatever ability is measured by performance on this task is being measured consistently throughout the experiment.

TABLE V

SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE PERFORMANCES ON THE THREE GROUPS

Trials	Diff.	r	σ diff.	C.R.	P.
No-Light minus Light-Right					
1– 3.....	1.57	.26	.70	2.27	99
4– 6.....	.30	.63	.32	.94	82
7–11.....	.13	.91	.29	.45	67
1–11.....	1.39	.67	1.34	.91	82
Light-Wrong minus Light-Right					
1– 3.....	.53	.29	.63	.84	79
4– 6.....	.19	.76	.39	.48	67
7–11.....	.03	.90	.35	.09	54
1–11.....	.72	.79	.95	.76	77
No-Light minus Light-Wrong					
1– 3.....	1.05	.43	.47	2.23	99
4– 6.....	.11	.64	.56	.20	58
7–11.....	.16	.83	.35	.46	67
1–11.....	.67	.70	1.16	.58	72

We have found, during the first three trials, a difference slightly greater than a chance expectation in favor of performance on the light-right and the light-wrong groups of

TABLE VI

RELIABILITY COEFFICIENTS

Group	Correlation between Odd and Even Trials	Reliability Coefficient of Entire Group (Spearman-Brown Formula)
Light-right.....	.83	.91
Light-wrong.....	.78	.88
No-light.....	.87	.93
All.....	.84	.91

syllables. The coefficients of correlation presented in Table VII indicate that even in these trials, where the largest differences between performances occur, the effect of the illumination is extremely variable. Illumination, apparently, does not effect the performances of all of the subjects in the same way. This supports our contention that whatever slight

effect there may be as a result of illumination, this influence is rather unreliable.

TABLE VII

INTERCORRELATIONS BETWEEN TOTAL NUMBERS OF ERRORS FOR THE THREE GROUPS

1 = light-right, 2 = light-wrong, 3 = no-light

Trials	r_{13}	r_{23}	r_{12}
1– 3	.29	.43	.26
4– 6	.76	.64	.63
7–11	.90	.83	.91
1–11	.79	.70	.67

We have criticized the results and interpretations of several investigators on the basis of their failure to control the factor of the informative cue character of the 'emphasis' or 'something happening after a response.' In the present experiment the illumination was not mentioned to the subjects until after they had learned the correct responses. There was an equal frequency of occurrence of the light-right and light-wrong conditions. These occurred in a different order for each trial. Thus it appears improbable that the illumination constituted an informative cue of correctness of choice.

In order to be sure that the illumination did not constitute a signal of success or failure the subjects were asked to answer the following questions immediately after the experiment:

1. Did you notice the flashing of the light?
2. Did you find it annoying ( ) indifferent ( ) pleasing ( )?
3. What meaning did you attach to it (as to the correctness of your response)?
4. Did you experience fatigue during the experiment? None ( ), Some ( ), Quite a little ( ), Very much ( ).
5. Did you repeat the correct syllables to yourself during the intervals between exposures?

All of the thirty-six subjects answered the first question in the affirmative. For the second question, eight answered

'annoying' and twenty-eight 'indifferent.' None of the 36 thought that the light was a signal of correct or incorrect choice. They answered that it had no meaning for them. The results of three other subjects were discarded on the basis of their answers to this question. One of these thought that the light meant an incorrect choice. Another thought that it meant 'right.' The third said that she "thought it was put there to annoy the subject." Nine of the subjects experienced no fatigue; twenty-one, some; five, quite a little; and one, very much. Only two did not repeat the correct syllables during the intervals between exposures.

On the basis of these reports we may assume that the light was not an informative cue of success or failure.

Conclusions.—Tolman, Hall and Bretnall's Law of Emphasis and Stephens' statement that "the mere fact that 'something' happens after a response tends to strengthen the underlying connection" are not substantiated by the results of the present investigation. If their hypotheses were true of learning in general, 'emphasis' would aid learning when it follows the correct responses. The data reported here show a slight, but statistically insignificant, tendency for this to occur. Their statements also imply that 'emphasis' following the incorrect responses would tend to stamp in those responses, thus affecting performance adversely. In the present experiment there was a slight tendency for the opposite effect to occur.

In view of the data presented here, and in the absence of experimental proof of the effect of 'emphasis' uncomplicated by the factor of informative cue, it is concluded that the retroactive effect of stimuli subsequent to a given response is due, when it does occur, not to 'emphasis' as such, but to the informative character of those stimuli.

(Manuscript received April 15, 1936)

REFERENCES

1. GLAZE, J. A., The associative value of nonsense syllables, *J. Genet. Psychol.*, 1928, **35**, 255–269.
2. GOODENOUGH, F. L., A note on Tolman's 'disproof' of Thorndike's Law of Effect, J. EXP. PSYCHOL., 1933, **16**, 459–462.

3. HALL, C. S., A repetition of Thorndike's experiments on improvement in the estimation of lengths, *J. Gen. Psychol.*, 1935, **9**, 238–241.
4. HULIN, W. S., AND KATZ, D., A comparison of emphasis upon right and upon wrong responses in learning, J. EXP. PSYCHOL., 1935, **18**, 638–642.
5. MUENZINGER, K. F., Motivation in learning. I. Electric shock for correct responses in the visual discrimination habit, *J. Compar. Psychol.*, 1934, **17**, 267–277.
6. MUENZINGER, K. F., Motivation in learning. II. The function of electric shock for right and wrong responses in human subjects, J. EXP. PSYCHOL., 1934, **17**, 439–448.
7. STEPHENS, J. M., The influence of punishment on learning, J. EXP. PSYCHOL., 1934, **17**, 536–555.
8. THORNDIKE, E. L., *The fundamentals of learning* (New York: Bureau of Publications, Teachers College, Columbia University, 1932).
9. TOLMAN, E. C., HALL, C. S., AND BRETNALL, E. P., A disproof of the law of effect and a substitution of the laws of emphasis, motivation, and disruption, J. EXP. PSYCHOL., 1932, **15**, 601–614.

ADDITIONAL LABORATORY EXPERIMENTS WITH QUALITATIVE WHOLES

BY M. V. SEAGOE

Earlier articles have pointed out that one cause of confusion regarding the whole-part problem is the failure to define the terms 'whole' and 'part' in a qualitative rather than a quantitative sense, and demonstrated that the superiority or inferiority of whole or part method varies roughly with the degree of integration of the unit selected. The present series of experiments tests the hypothesis with added types of material meeting the same criterion under a slightly modified procedure with a more carefully controlled group of subjects.

The subjects were thirty-four boys within one year of twelve chronologically paired according to intelligence quotient. A comparison of the two groups on these factors shows a statistically insignificant difference in favor of Group I consisting of 0.24 months in age, 0.82 points in intelligence quotient, and 1.24 months in mental age. In no case does the chance of a true significance lie above 57. To further control this factor, the method of presentation was rotated for each subject, and coefficients of correlation between mental age and efficiency in mastering each problem were computed. For mirror drawing, number code, block design, and chess pattern the correlations with time scores ran $-.051 \pm .115$, $.443 \pm .093$, $.302 \pm .105$, and $.282 \pm .106$, respectively; similarly for error scores the relationships were $.023 \pm .116$, $.352 \pm .101$, $.169 \pm .112$ and $.219 \pm .110$.

Group I learned the mirror drawing problem by the whole method, the number code by the part method, the block design by the part method, and the chess pattern by the whole method. Group II used the opposite method with each problem.

The time of day for original learning was held constant, all experiments being completed between eight and eleven A.M. and both members of the experimental pair coming on the same day with a rotated order for priority of appointment. Retention was checked in the same way twenty-eight days later. Participation in the experiments was voluntary and it was freely given. There was no warning at the time of learning that retention would be checked and no opportunity for practice between learning and retention check.

Experiment 1. Mirror Drawing

The first problem presented to the subject was to trace a diamond-shaped groove in a brass plate with only the mirrored image of the hand and pattern visible. The outside

measurements of the groove at the two axes were four and one-half and six and one-half inches, the groove being one-fourth inch wide with notches three-sixteenths of an inch wide and one-eight inch deep every three-sixteenths of an inch along both sides to prevent leaning. The bottom of the groove was lined with tin insulated from the plates above and below by cardboard so that when the sides of the groove were touched an electrical circuit was completed, and erorrs recorded on an impulse counter. There was also a cover plate which concealed the rest of the figure from the subject while one side was being presented in the part method.

The reliability of the apparatus was estimated in two ways. Using odd-even trials with the first omitted and the Spearman-Brown formula coefficients of .970 ± .007 for time and .995 ± .001 for errors were obtained. Such coefficients are spuriously high because of the lack of independence of the two series, because mere increment during learning tends to produce correlation independent of the reliability of the material, and because of the artificial constancy of certain factors when the Spearman-Brown formula is used. In the second place, rho was computed for the seventeen cases learning by the whole method and r estimated from it. Intercorrelations between the three segments of the curve on time and error scores varied from .57 to .93, the most reliable figures approaching .87.

The subject was aided in placing the stylus at the starting point, the direction of tracing indicated, and record made of time and errors. The first five trials constituted the period of presentation by differing method, trials six through ten the period of checking mastery of the entire problem using the whole method in all cases, and trials eleven through fifteen the retention check made twenty-eight days later. Whole presentation for the first five trials consisted of tracing the entire figure five times; part presentation consisted of five tracings each (in the direction used in whole tracing) of sides one, three, two, and four, only the side traced being visible during the process. The retention check involved the same procedure as trials six through ten.

Learning curves are presented in Figs. 1 and 2. The solid line represents results for the whole method, the broken line for the part method in each of the figures. In terms of time there is little difference, whole presentation having a slight advantage. In terms of errors part presentation is superior.

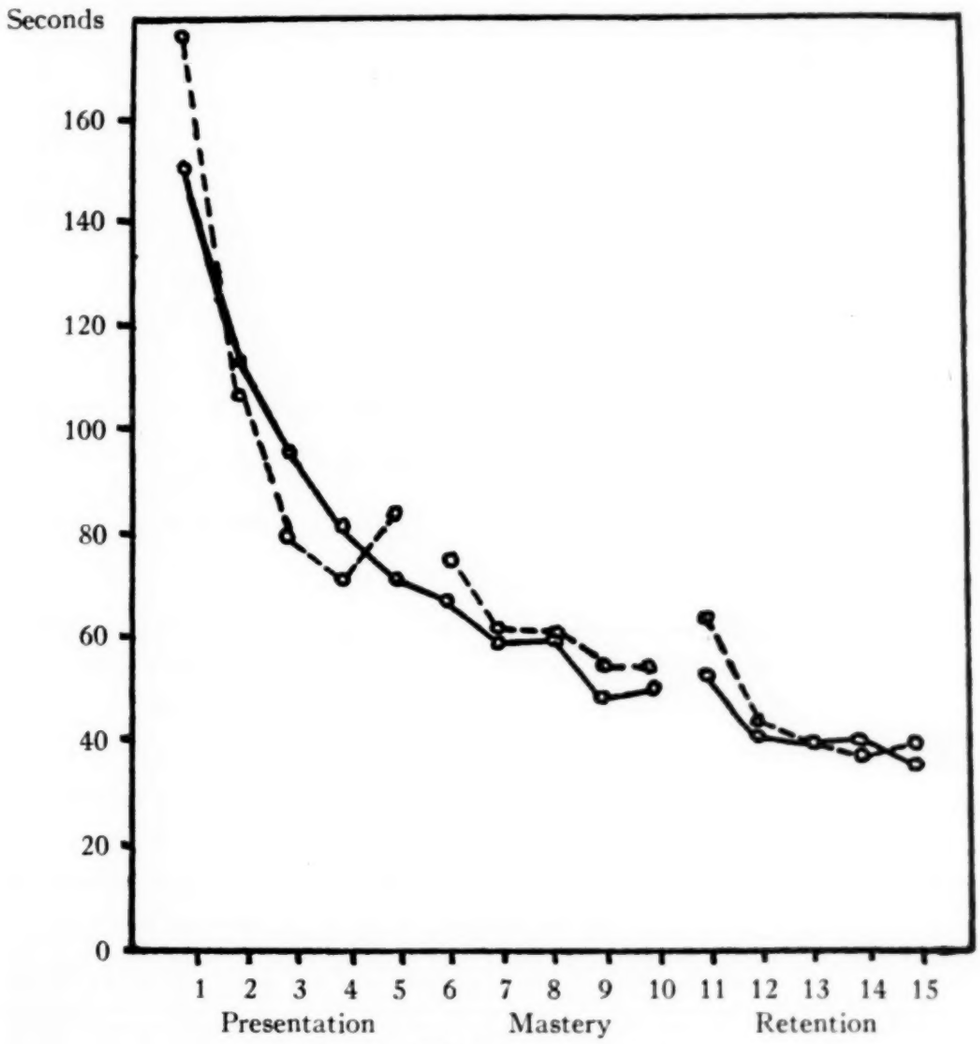

FIG. 1. Mirror drawing: time.

These tendencies are substantiated by the statistical summary given in Table I.

Such contradictory results present a dilemma; generalization seems impossible. Fortunately, later experiments yielded more consistent results. The mirror-drawing failed to conform to the pattern set by the number code, block design, and chess pattern for two possible reasons. In the

first place, mirror drawing has a larger component of skill and less of ideation than the others. If the difference between the methods involves perception of relationships, the difference would tend to be greater on problems involving more ideation and less motor control. In the second place, there

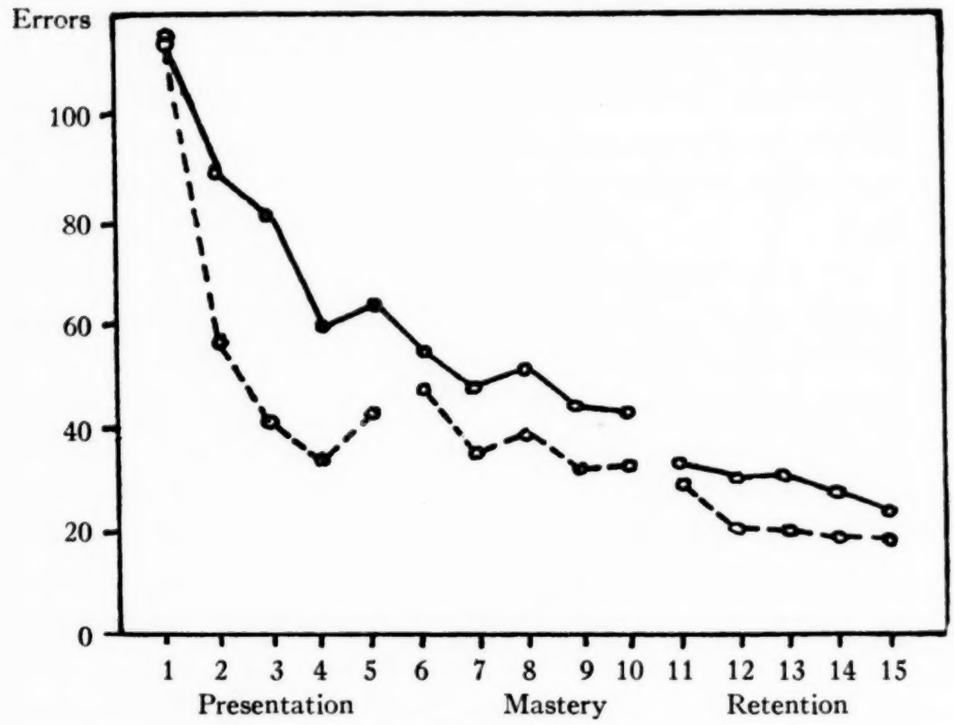

FIG. 2. Mirror drawing: errors.

TABLE I.

RESULTS, EXPERIMENT I

Item	Time on Trials				Errors on Trials			
	1–5	6–10	11–15	Total	1–5	6–10	11–15	Total
Mean: Whole method...	511.59	287.88	213.47	1012.94	409.65	241.65	146.29	797.59
Part method.....	518.00	308.59	227.18	1053.76	290.29	186.53	108.76	585.59
P.E. mean: Whole.....	42.47	22.32	11.80	69.88	35.54	20.63	15.10	65.78
Part method.	34.39	16.90	12.10	56.19	36.69	19.34	13.16	65.26
Difference in means.....	W6.41	W20.71	W13.71	W40.82	P119.36	P55.12	P37.53	P212.00
P.E. of difference.....	54.65	28.00	16.90	89.67	51.08	28.28	20.03	92.66
Critical ratio.....	W0.12	W 0.74	W 0.81	W 0.45	P 2.34	P 1.95	P 1.87	P 2.29
Probability reliable.....	W53	W69	W71	W62	P 94	P91	P89	P 94

is less difference between the whole and its segments as well as between segments in the mirror drawing than in the other problems; segments are the same except for orientation. It is perhaps questionable whether the problem involves a true qualitative whole.

The results of Experiment 1 are inconclusive, possibly because of the nature of the material used.

Experiment 2. Number Code

The plan of the number code consisted of two horizontal lines intersected by two vertical lines, the numbers 1 through 9 being written in the spaces thus formed. The Fig. 1 was to be represented by the lines surrounding it, *i.e.* a right angle opening up and to the left, Fig. 2 by three sides of a square opening up, 3 by a right angle opening up and to the right, 4 by three sides of a square opening to the left, 5 by a complete square, 6 by three sides of a square opening to the right, 7 by a right angle opening to the left and down, 8 by three sides of a square opening down, and 9 by a right angle opening to the right and down. Practice forms were devised consisting of a practice section and a testing section presented in a cardboard frame so that the code was visible only during practice.

The reliability of the code was also estimated in two ways. On odd-even trials the coefficients were .963 ± .008 for time scores and .964 ± .008 for errors. Estimating r from rho the intercorrelations between segments of the curve range from .84 to .93.

The subject was instructed that he was to learn to write the form enclosing the number for the number and presented the practice material. At the end of thirty seconds in the group using the whole method or of ten seconds for the part method practice was stopped, and the subject asked to write in as many substitutions as possible on a test sheet. Time required to write the figures and number of omitted and incorrect responses were recorded. The procedure was repeated until five trials were completed.

The group which learned by the whole method worked from the entire code presented in pattern from beginning to end. The group which learned by the part method was given five trials on the symbols for 1, 4, and 7 written serially, five on 2, 5, and 8, and five on 3, 6, and 9. Both groups worked with the entire code from trial six on.

Learning curves are presented in Figs. 3 and 4. Both time

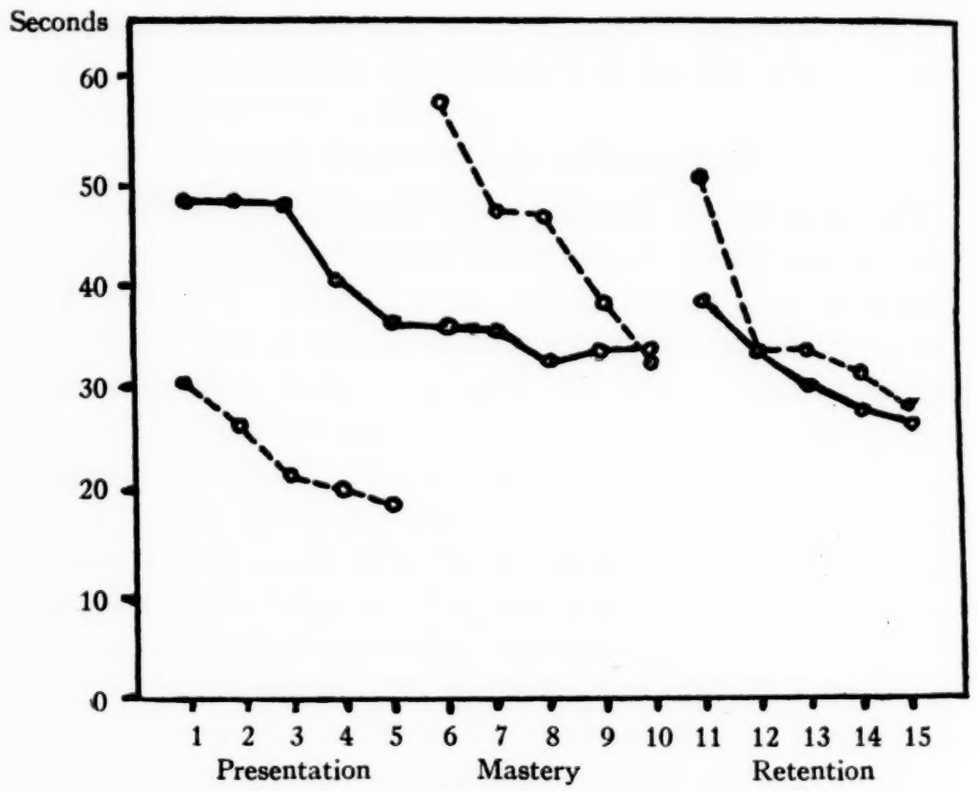

FIG. 3. Number code: time.

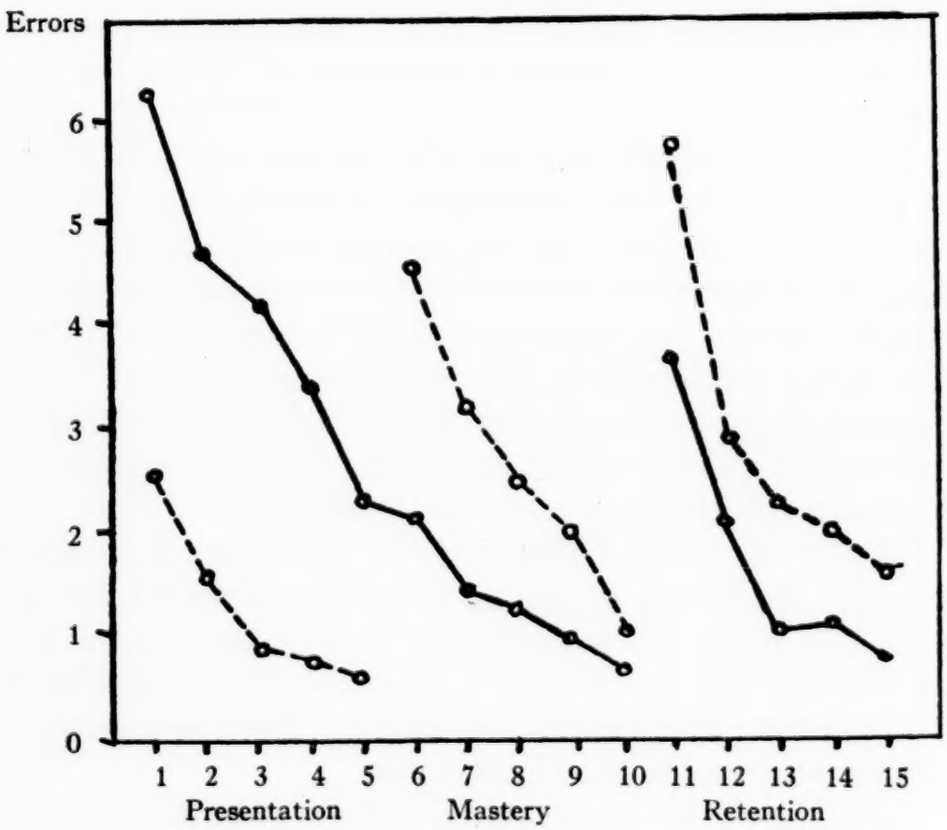

FIG. 4. Number code: errors.

and error curves show part presentation more economical on the first five trials, whole presentation superior on tests of mastery of the entire Figure and of retention. That is, part presentation results in a saving in time and errors during actual initial exposure, but the learning does not carry over to mastery of the entire code. The difference is statistically small but consistent.

The tendency of the lines representing the two methods to converge from trial six on may be attributed to an increasing amount of sameness of practice beginning at that point.

The greater economy of whole presentation in mastery and retention is present in spite of the fact that the method of practice and testing made attention to parts or number symbols rather than to the whole number pattern the procedure for learning. This fact should theoretically give the part method the advantage from the point of view of transfer, yet apparently the factor of whole presentation more than compensated.

Table II presents a statistical summary by segments. The fallacy of combining data from the entire curve is clearly

TABLE II

RESULTS, EXPERIMENT 2

Item	Time on Trials				Errors on Trials			
	1–5	6–10	11–15	Total	1–5	6–10	11–15	Total
Mean: Whole method...	224.82	173.53	156.41	554.76	20.94	6.71	8.94	36.59
Part method....	121.35	222.94	178.18	522.47	6.53	13.53	14.65	34.71
P.E. mean: Whole.....	9.55	11.72	8.42	25.85	1.75	1.33	1.61	4.19
Part method	6.33	18.74	8.98	30.44	0.69	1.60	1.65	3.58
Difference in means.....	P103.47	W49.41	W21.77	P32.29	P 14.41	W 6.82	W 5.71	P 1.88
P.E. of difference.....	11.46	22.10	12.31	39.04	1.88	2.08	2.31	5.51
Critical ratio.....	P 9.03	W 2.24	W 1.77	P 0.81	P 7.66	W 3.28	W 2.47	P 0.34
Probability reliable.....	P100	W93	W88	P71	P100	W99	W95	P59

demonstrated in the fact that fairly reliable differences in opposite directions are found for initial presentation on the one hand and mastery and retention checks on the other. Combination of all trials, however, yields a comparatively unreliable difference.

Experiment 2 shows that, although the part method results in an immediate saving of time in the process of pres-

entation, the whole method is more economical in achieving mastery of the entire material and in its retention.

Experiment 3. Block Design

The design used was the same as problem 4 in the earlier series of experiments, and the procedure was comparable.

The reliability of the design with the present group of subjects computed from odd-even trials was .942 ± .013 for time scores and .892 ± .024 for errors. Estimating r from rho the intercorrelations between segments of the curve range from .12 on retention with test of mastery to .83 on presentation with retention. The true reliability probably lies between .70 and .85.

The instruction was the same as for the adults except in the case of a few subjects who found solution very difficult, under which circumstances "Do you see this design? Can't you find the blocks that make it?" was added. The procedure differed in that five trials were given to all subjects by varying method, five for checking mastery by the whole method, and later five for checking retention by the whole method in place of learning to mastery.

The group learning by the whole method worked with the total design from the beginning. The group learning by the part method had five trials each with the four blocks in the upper left corner followed by similar presentation of those in the lower right, upper right, and lower left.

Learning curves for Experiment 3 are presented in Figs. 5 and 6. They are essentially the same as those of Experiment 2, with a greater weighting in favor of the whole method. Subjective observation showed that subjects who had originally learned by the part method almost never followed the parts originally learned in their solution of the total design, but followed the pattern instead; only one case in trials six through ten and none in the retention check deviated from this procedure.

Table III presents a statistical summary by segments. The part method is again somewhat more economical in original presentation, but the whole method is more economi-

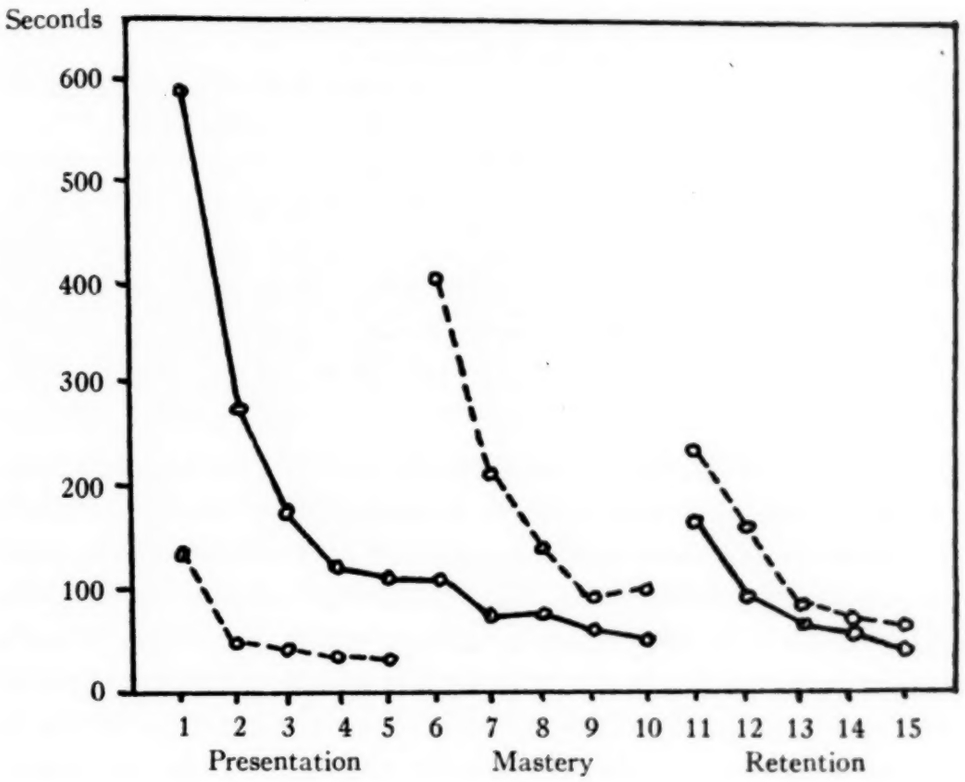

FIG. 5. Block design: time.

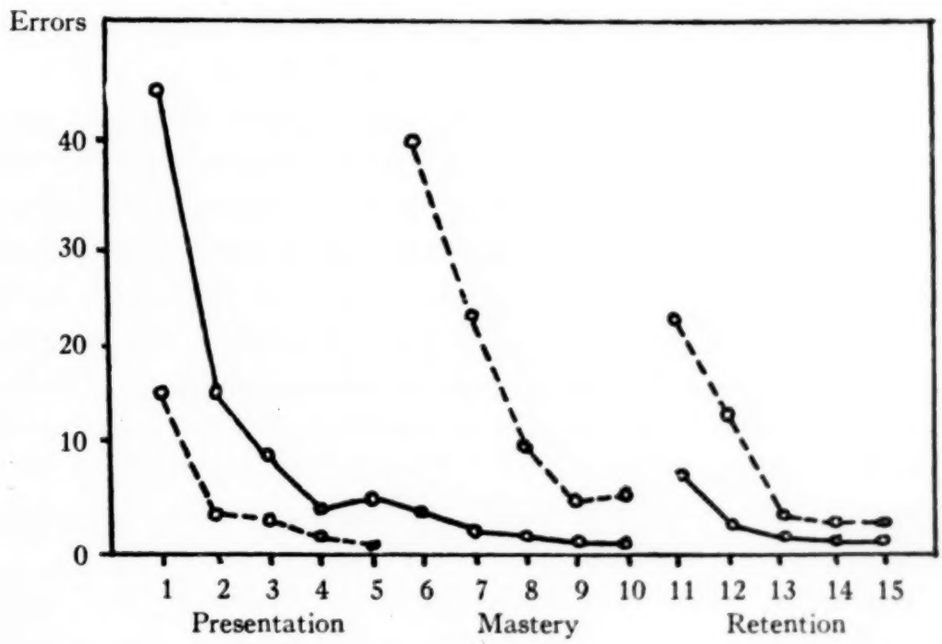

FIG. 6. Block design: errors.

TABLE III

RESULTS, EXPERIMENT 3

Item	Time on Trials				Errors on Trials			
	1–5	6–10	11–15	Total	1–5	6–10	11–15	Total
Mean: Whole method.	1293.76	387.00	441.88	2122.65	82.35	10.82	13.53	106.71
Part method...	306.41	944.06	610.24	1860.71	26.24	84.65	44.71	155.59
P.E. mean: Whole....	383.76	34.77	25.11	114.48	7.31	2.12	1.58	8.87
Part.....	22.72	77.68	61.93	125.98	3.19	8.05	5.92	13.72
Difference in means...	P987.35	W557.06	W168.36	P261.94	P 56.11	W 73.83	W 31.18	W48.88
P.E. of difference.....	384.43	85.11	66.83	170.23	7.98	8.33	6.13	16.34
Critical ratio.....	P 2.57	W 6.55	W 2.52	P 1.54	P 7.03	W 8.86	W 5.09	W 2.99
Probability reliable...	P 96	W100	W 95	P 85	P100	W100	W100	W98

cal in terms of mastery and retention. This particular experiment perhaps brings out the concept most clearly. There is a closed, integrated figure, presented as a whole and broken into arbitrary segments. The tendency of the learner to follow the form in his solution, a form which is hidden in part presentation, results in the superiority of the whole method as judged by mastery and retention. Where, therefore, there is a high degree of unity the whole method of presentation seems more economical. This generalization might not hold for loosely integrated or non-ideational materials.

Experiment 3, in summary, shows the same tendencies as Experiment 2.

EXPERIMENT 4. CHESS PATTERN

The material consisted of six chess figures arranged in a pattern on a chess board. The figures, which were all the same color, formed a triangle, the pawn at the point in front, the king, knight and queen along the line at the back, and the castle and bishop standing on the diagonals from the pawn to the king and queen respectively. Unity of the pattern was obtained both by the triangular arrangement and by the progression from larger figures at the back to the smaller ones of similar design at the front. The figures were placed in a constant position on the board, a screen covering the entire board during placement.

The reliability of the chess pattern computed from odd-even trials was .942 ± .013 for time scores and .892 ± .024 for errors. Estimating r from rho the intercorrelations

between segments of the curve range rather widely, possibly because of defects in the scoring scheme, but the most reliable figures lie between .67 and .89. Because of the shorter time involved in learning and the difficulty of scoring, the results of Experiment 4 are of value only in supplementing other findings.

The subject was seated before the screen, instructed that he would be shown some figures for two or six seconds, depending on the method of learning, then asked to replace the figures. Time and errors in reconstructing the pattern were recorded. The amount of error was computed by counting the number of moves distant from the correct position each figure was placed and totaling the result. This method has the disadvantage that, although the figures may have been placed in their proper relationships, if the pattern was wrongly placed on the board the error score was comparatively large, while if he failed to perceive the relationship yet placed them near the proper place on the board his error score might be only moderately large. This factor should theoretically decrease the superiority of whole presentation; at any rate it was constant for both methods.

The group learning by the whole method worked with the entire set from the beginning. The group learning by the part method had five trials each on the queen and bishop, king and castle, and knight and pawn. From trial six on the whole procedure was used with both groups.

Learning curves are presented in Figs. 7 and 8, and a statistical summary by segments in Table IV. The findings

TABLE IV

RESULTS, EXPERIMENT 4

Item	Time on Trials				Errors on Trials			
	1–5	6–10	11–15	Total	1–5	6–10	11–15	Total
Mean: Whole method	65.00	44.88	62.18	172.06	14.82	4.71	10.88	30.41
Part method	37.47	74.06	67.82	179.35	3.00	17.88	14.29	35.18
P.E. mean: Whole	4.27	2.11	3.55	8.33	1.51	0.98	1.96	2.69
Part method	0.90	3.96	3.58	6.56	0.89	2.13	1.75	3.56
Difference in means	P 27.53	W 29.18	W 5.64	W 7.29	P 11.82	W 13.17	W 3.41	W 4.77
P.E. of difference	4.36	4.49	5.04	10.60	1.75	2.35	2.63	4.46
Critical ratio	P 6.31	W 6.50	W 1.12	W 0.69	P 6.75	W 5.60	W 1.30	W 1.07
Probability reliable	P100	W100	W76	W68	P100	W100	W81	W76

are of the same nature as those obtained with the number code and the block design, that is the part method is economi-

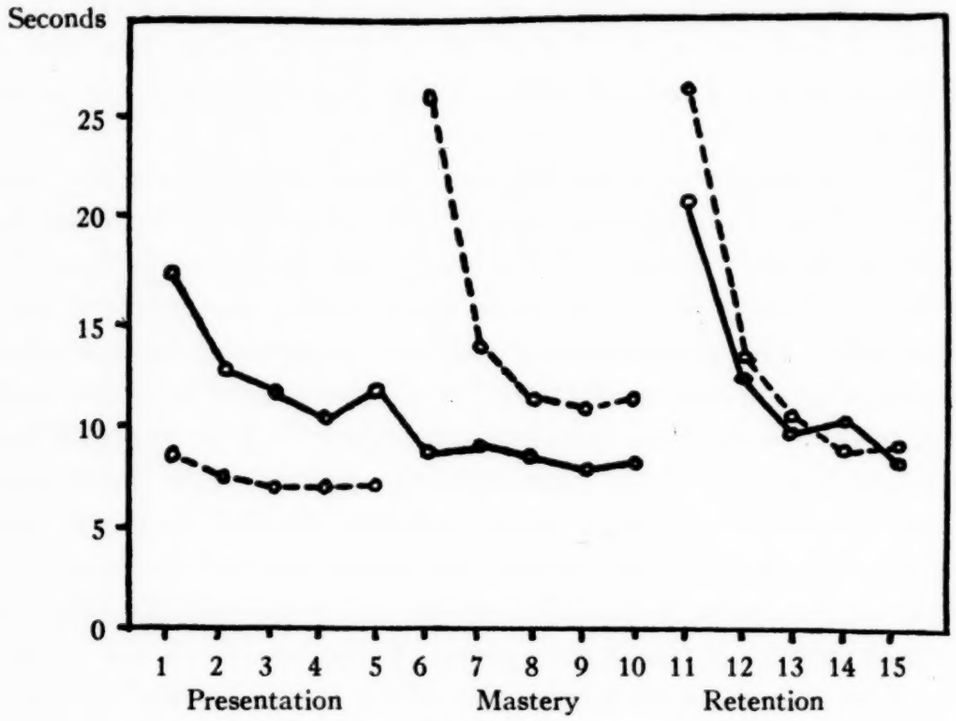

FIG. 7. Chess pattern: time.

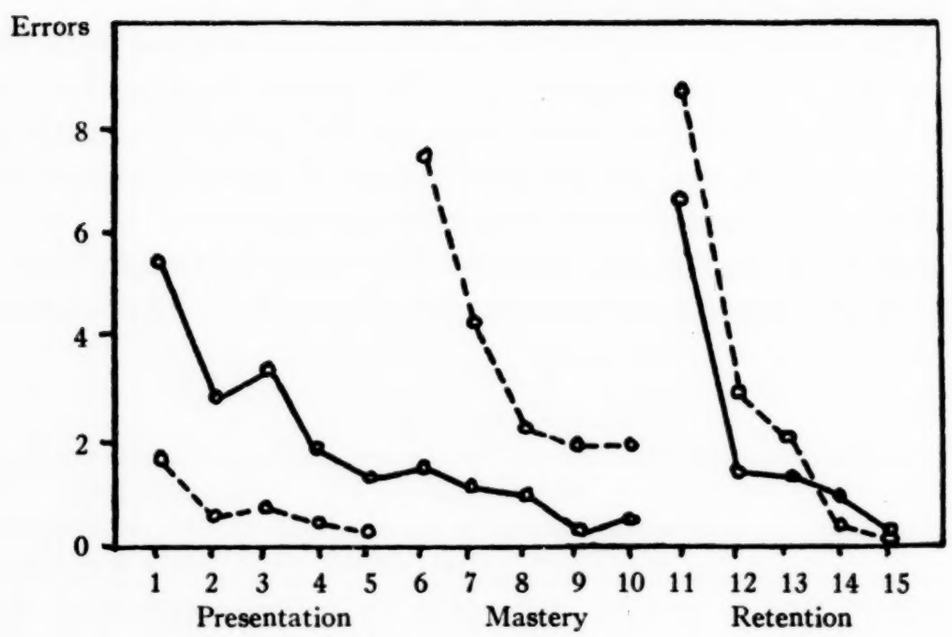

FIG. 8. Chess pattern: errors.

cal in original presentation, the whole method in mastery and retention.

Summary and Conclusion

Intercorrelations between the problems based on time scores are presented in Table V. They are comparatively low, indicating that the materials actually represented differing applications of the concept of wholeness in the unit for learning. The table brings out the difference between the mirror-drawing problem and the other materials discussed earlier.

The results for the four experiments seem to justify certain generalizations. (1) The part method results in a saving in time and errors in original presentation.[1] (2) The whole method is most economical in terms of mastery. (3) The whole method is most economical in retention. (4) Pooling

TABLE V

Intercorrelations

Problem	Number code		Block design		Chess pattern	
	Time	Errors	Time	Errors	Time	Errors
Mirror drawing...	−.086±.115	.096±.115	.198±.111	.440±.093	.087±.115	.334±.103
Number code....			.120±.114	−.104±.114	.559±.080	.045±.115
Block design.....					.228±.106	.379±.099

the results on all trials results in no clear differentiation between the methods. (5) These generalizations apply to the problems involving a large ideational component, but the problem involving motor skill for the most part fails to show much difference between the methods. (6) The differences supporting these generalizations, though small, are consistent. The absolute size of the individual ratios is probably not significant, but the consistency of the results argues for their validity.

A word of explanation in regard to the decreasing difference between the two methods as learning progresses is needed. The group learning by the part method used a different procedure only on trials one through five; thus part

[1] The designation of trials 1 to 5, 6 to 10, and 11 to 15 as presentation by differing method, mastery, and retention respectively has been introduced to help keep in mind the variations in procedure represented in these segments; the terms are used in a specific sense somewhat different from the usual connotation.

learning included some whole learning while whole learning included no part learning. The actual difference between the methods, therefore, is most clearly indicated at trial six rather than by an average of succeeding trials, and is greater than that shown by the data. The tendency for the difference to be less in retention than in mastery may be a function of the increasing number of trials by the whole method in the part group, rather than of any actual tendency for the difference to drop out during an interval.

In conclusion, when a whole is defined as a Gestalt with important inner relationships, and when that unit involves a relatively large ideational factor, the material is more economically presented as a unit rather than as segments as judged by efficiency of mastery and by retention. Part presentation, however, saves time in the process of presentation, although mastery of the parts does not assure mastery of the whole.

(Manuscript received January 10, 1936)

AN EXPERIMENTAL ISOLATION OF SOME FACTORS DETERMINING RESPONSE TO RHYTHMIC CUTANEOUS STIMULATION. II. TEMPERATURE

BY ROGER M. BELLOWS

Division of Standards and Research, United States Employment Service

In a former paper [1] the influence of the variables of frequency, pressure, and time of intermittent stimulation upon the cutaneous flicker limen was reported. Two other controllable factors are involved as determiners of response to rhythmic skin stimulation. These are: (1) the temperature of the stimulus; and (2) skin adaptation (acclimatization) to temperature.

Apparatus and Procedure

Allen (1) and his students first devised the air blast technique of intermittent stimulation for the investigation of vanishing flicker. Apparatus somewhat similar to that used by Allen was employed here for the study of the variables of frequency, pressure, and time. It was described in the former paper by the present writer. The apparatus for the former study consisted of an air source and manometer for the purpose of controlling and measuring pressure; a sectored disc to produce intermittent stimuli; and a recording system (polygraph) for measuring frequency of successive stimuli and their duration. Slight modifications of this apparatus were made for the investigation of the temperature variable.

Allen did not isolate the variable of temperature of his intermittent stimuli. He experimented with different methods of rapid successive thermal stimulation. These workers tried warmed air by passing it through a heated copper coil. They report, "The sensation of warmth was readily evoked, but even an air temperature as high as 80° C. failed to yield a flickering sensation with a motor speed as low as one revolution per second" (2). The pressure of the stimulus was maintained at a very low intensity, so that the temperature influence might be distinguished from the pressure variation and the attention concentrated on the former.

Allen did not note and explicitly measure the influence of temperature of the air pressure stimuli upon vanishing flicker frequency, he assuming that the two sensations were subtended by diverse receptors and qualitatively different nerve fibers.

Various methods of intermittent thermal stimulation were tried by these workers. Allen decided upon electrical stimulation—"Attention was then directed to the possi-

[1] Both this and the former investigation (3) were conducted at the Ohio State University. The writer is indebted to Professor Robert D. Williams of the Ohio State University for his helpful guidance.

bility of producing stimulations of the thermal receptors by an intermittent electric current. After various trials some success was obtained in eliciting sensations of warmth, cold, pressure, and pain" (2).

We were interested in the influence of the temperature of the intermittent stimulus upon frequency for vanishing flicker when the other controllable variables were maintained constant.

Several minor additions to the apparatus and modifications of the method were necessary. The air hose from the storage tank was cut and the air directed into a two liter hard glass flask which was heat insulated by wrapping with asbestos and which was placed (for the warm temperatures) over an electric hot plate. The air was forced from the flask into a temperature control chamber in which was inserted a hard glass tube shaped as a double "U," one end of which was closed and sealed into the control chamber. This was filled with toluene, and a capillary tube filled with mercury was attached to the open end. Upon expansion of the toluene (caused by heating the air in the chamber) the mercury would rise in the tube, making contact with the top and closing a circuit leading through a relay and a 6 volt storage battery. When the mercury made contact, thereby closing the circuit, a relay magnet was energized. The relay cut off the electric hot plate. Heat was maintained within less than 1° C. of the desired temperature and adjusted by raising or lowering the top contact in the thermostat. The air temperature was measured 5 cm back of the air nozzle by a thermometer sealed in a glass container through which the air was forced. This measurement was called 'air temperature No. 1.' The air was then measured 1 cm in front of the nozzle, the locus of the stimulated area, and the change calibrated. For cooling the air dry ice was used.

Some difficulties were encountered in calibrating the change in temperature of the air during each experimental period, just as it struck the skin of the subject. This was not feasible. The next best procedure for measurement of the temperature of the air stimulus would be to accurately calibrate the change from the measured temperature back of the nozzle. The best thermometer available had a mercury bulb of a diameter of approximately .3 mm. The inside diameter of the air nozzle was .1 mm. The measurement 5 cm back of the nozzle was called measurement No. 1. Measurement was made, for calibration, 1 cm in front (toward the subject's position) of the nozzle and was called measurement No. 2. Thermometer No. 1 was marked in 1° C. divisions but they could be read with a fair degree of accuracy, with care, in .5° units. Finer units were estimated. The scale on thermometer No. 2 was divided into units of .2° C. During the calibration the room temperature was 23° C. The equivalents are given in Table I.

There was one limitation to this method of calibration. Thermometer No. 2 was pointed, mercury bulb foremost, toward the opening of the nozzle. The heated (or cooled) air struck the somewhat rounded end of the mercury bulb and effused into the surrounding air. The air in contact with the sides of the thermometer was more nearly of room temperature. Hence the readings on thermometer No. 2 are probably somewhat lower than the actual temperature of the experimental air as it struck the cutaneous surface of the observer.

We were more interested, however, in the relative influence of temperatures than its absolute measurement. For purposes of repeating the experiment the temperature measurement method we believe has been adequately described.

In the temperature experiments reported below the stimulus temperature is always reported as of temperature reading No. 1 unless otherwise specified.

TABLE I

SHOWING CALIBRATION OF TEMPERATURE, READING 1 AND READING 2, EQUIVALENTS

Reading 1, °C.	Reading 2, °C.
0	7
10	14
18	19
23	23
34	26.6
35	27
36	27.6
37	28
38	28.6
39	29
40	29.2
45	32
50	34.6
55	38
60	41
65	45
70	49
75	53
80	57

In the temperature experiment the air temperature and air pressure were set and recorded, room temperature recorded, and the O seated with head in head rest, and was instructed to take readings. All readings for temperature were taken at intervals of two minutes to prevent adaptation effects or intangible tuning influences, and with rest pauses after each five trials.[2]

Method for Determining the Influence of Temperature Adaptation upon Vanishing Flicker Frequency.—After it was discovered in the preliminary experiment that the temperature of the air blast was apparently a factor in response to fusion of successive stimuli it was decided to attempt to determine any possible influence of an adaptation temperature. It was thought that, if the skin were adapted to a temperature of 37° C. for example, this might influence the frequency.

Two conditions of skin adaptation were chosen. In condition I the skin was adapted to 37° C. In condition II the skin was adapted to 18° C. Under both conditions of adaptation the stimulating air was 37° C. Four Os were used, C, Fi, Fa, and G. All of these had participated in the temperature experiments, and were trained subjects.

When the subject came into the experimental room he was seated before the apparatus, and the lips, the experimental area used, coated with vaseline for the purpose of preventing drying and to mitigate the possible influence of humidity of the air stimulus. The experimenter heated or cooled water in a flask to the desired temperature, either 37° C. or 18° C. Cotton, wrapped around the bulb end of a thermometer, was saturated with the water, then pressed firmly against the cutaneous area of the subject which was to be stimulated, and held there for two minutes, at which time it was again saturated with water of the desired temperature and immediately

[2] Os were instructed to present the most sensitive surface or spot of the lower lip to the stimulating air jet by slight adjustive movements of the head. The work of Earhart, H. and Dallenbach, K. M. (5) indicates that "intervals of one to five minutes are sufficiently long for reliable and consistent results" when the more intense warm spots are stimulated. Similar intervals, for practical purposes, were found for the cold spots (4).

again applied to the skin for another two minutes. The temperature of the water saturated cotton as measured by the thermometer varied approximately 1.5° C. The subject was adapted for four minutes at the desired adaptation temperature. The area was then dried quickly with a cloth to prevent evaporation and cooling and the subject immediately took a reading of vanishing flicker frequency. The temporal interval between removal of the cotton and the click of the apparatus recording the reading varied from 20 to 25 seconds.

Results

The Temperature Variable.—The data presented schematically in Fig. 1 show the influence of temperature of the air stimulus upon fusion frequency. All of these data were taken at the pressure of 1.5 cm.

We were interested in determining the influence, if any, of temperature. Renshaw (**5**) found that temperature of solutions of low percentages of salt content influenced the absolute threshold of the subjects' reaction to salt. It was this fact that suggested the present problem to the writer.

It was found, under the experimental conditions of this investigation, that temperature of the air stimulus is apparently an important factor in vanishing flicker frequency; that, provided the experimental technique here used is sound, no experiment dealing with cutaneous phenomena and using the air blast method (*e.g.* Allen, Adrian, McKeen Cattell, and Hoagland) may well neglect this variable.

Four of the five subjects used in this part of the study were junior or senior undergraduate students with more than one year's work in psychology. The fifth was a graduate student in psychology.

The data are presented diagrammatically in Fig. 1. For subject C at 0° C. the vanishing flicker frequency was found to be 50, at 10° C., 59, etc. At room temperature, 23° C., subject G's average reading was 100 ± 4. This point is not clearly shown in the graph.

The findings indicate that for all subjects tried at the lower temperatures (subjects Fi and Fa did not take readings below room temperature) of 0° C. and 10° C. the vanishing flicker frequency was considerably lower than at room temperature; lower also than that found at temperatures between 50° C. and 55° C.

All of the curves show a decided and uniform tendency to decrease in frequency of fusion between the temperatures of

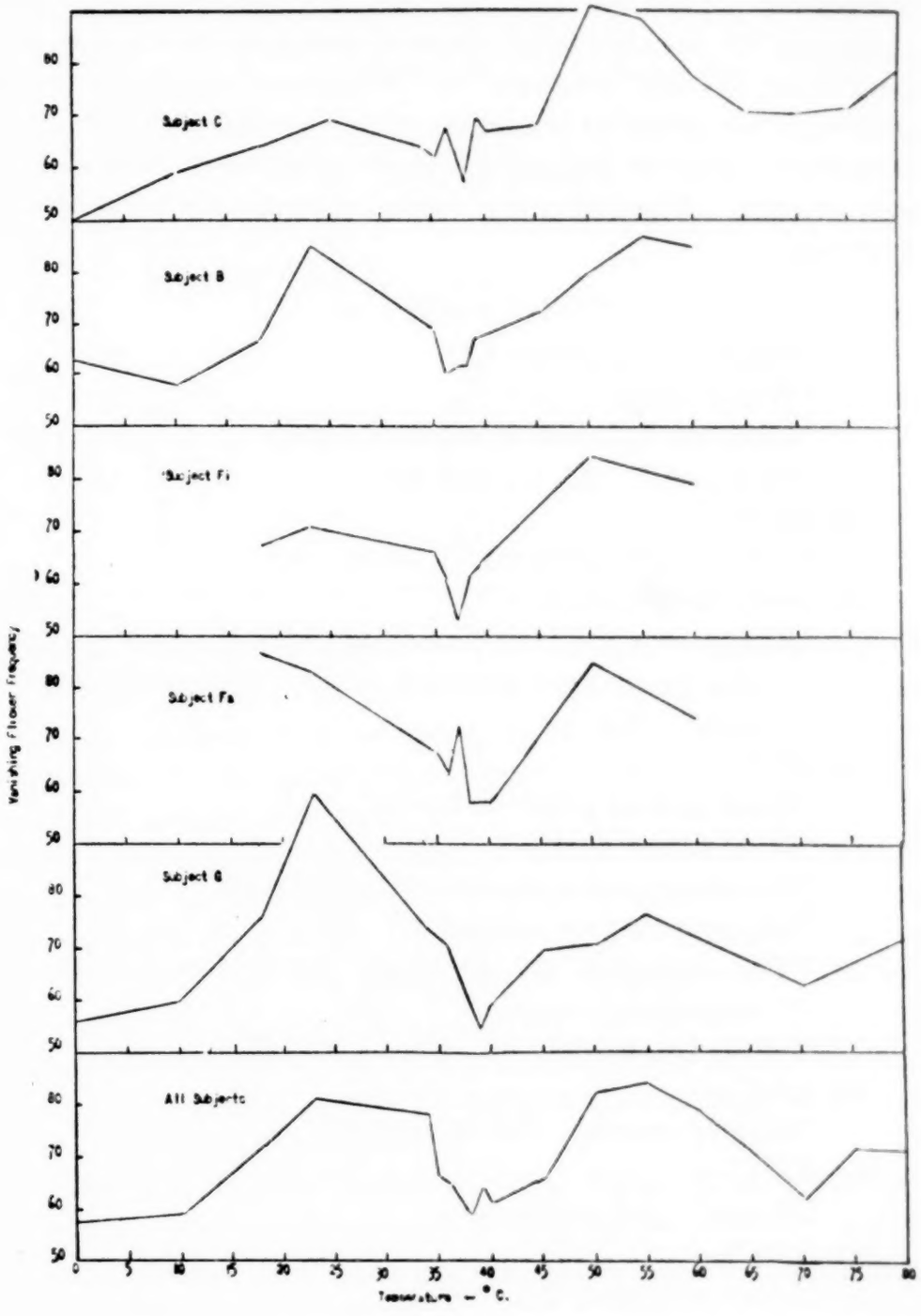

FIG. 1.

35° C. and 40° C. (measurement no. 1, see calibrated equivalents, Table I).

The verbal reports of the different subjects agree in each case as to: (1) the sureness of their phenomenal criterion of vanishing flicker for the different temperatures; (2) their judgment of the apparent relative pressure at various temperatures; (3) the 'coolness' or 'warmness' qualitative component of the stimulus under the varied conditions; and (4) the frequency range of flickering before 'continuous stimulation' was certain. The following verbal reports are presented as typical:

At 0° C.—
'Smooth' (subject C)
'Wider range' (C)
'Difficult to locate the point' (G)
'Very cool' (C, G, and B)
At 10° C.—
'Cool' (all subjects)
At room temperature—
'Range of vanishing flicker appears shorter—perhaps due to greater practice at this stimulus temperature' (G)
At 38° C.—
'Hard to find point of fusion' (G)
'Wide range' (C)
'No temperature component perceived' (C)
'No temperature sensation' (G)
"The impulses, at definitely lower frequencies, are smoother, 'rounder.'" (B)
'Lower intensity' (C, Fa, G, and B)
At 40° C.—
'Slightly warm' (all subjects)
At 60° C.—
'Warm' (all subjects)
At 80° C.—
'More compact, hot' (C)
'Heavier and hot' (G)

Adaptation to Temperature.—The data presented in Table II show a possible influence of adaptation temperature upon

TABLE II

SHOWING INFLUENCE OF TEMPERATURE AT WHICH CUTANEOUS SURFACE IS ADAPTED. STIMULUS INTENSITY 1.5 CM, STIMULUS TEMPERATURE 37° C. (AT SKIN)

Subject	Condition I Adapt. Temp. 37° C.	Condition II Adapt. Temp. 18° C.
C		
N	5.0	5.0
M	54.4	91.4
	$t = 5.35$ (significant difference)	
Fi		
N	3.0	4.0
M	70.0	67.2
	$t = .02$ (difference not significant)	
Fa		
N	5.0	5.0
M	58.2	91.6
	$t = 5.31$ (significant difference)	
G		
N	2.0	5.0
M	59.0	54.0
	$t = .473$ (insignificant difference)	

vanishing flicker frequency. The method for adapting the skin was summarized above. Under condition I the cutaneous surface was adapted to a 37° temperature. In condition II the adaptation temperature was 18° C. The stimulus temperature was 37° C. at the skin surface in all cases (temperature measurement no. 2). Subject C reacted to the first condition with an average frequency of 54; to the second with an average frequency of 91. A shift from one to another of the two conditions of temperature in the case of subject Fi did not effect a marked difference. The difference such as exists, between 70 and 67, is in the opposite direction from that shown by subject C, and is not a statistically significant difference.[3]

One might infer however, from these data, that some subjects tend to make use of (or perhaps be confused by) the

[3] The following formula which was rearranged for computational purposes by Harold A. Edgerton (of the Occupational Research Program, and is to be reported in his paper, a table to facilitate the computation of '*t*' for the comparison of means) was used:

$$t = \frac{M_1 - M_2}{\sqrt{N_1\sigma_1^2 + N_2\sigma_2^2}} \sqrt{N_1N_2 - \frac{2\,N_1N_2}{N_1 + N_2}}\,.$$

temperature component of the stimulus more than do others, in their perception of vanishing flicker frequency.

When the tabulations for all subjects are taken together the differences in favor of higher frequencies under condition II appear reliable, but here it is obvious that the readings of two of the four subjects, viz. C and Fa, bring about the difference. No inferences can be drawn other than the possible one mentioned above.

Summary

In addition to the effect of the factors of intensity and time upon vanishing flicker frequency the factor of temperature of the air stimulus was found to be a variable. It appears evident that for the lower temperature of 0° to 10° C. (measured back of the air nozzle) the vanishing flicker frequency was lower than at room temperature, other conditions of time and pressure being maintained constant. Pressure was maintained at 1.5 cm mercury. Stimulus temperatures of 35° to 45° C. also resulted in lower vanishing flicker frequencies on all of the five subjects who participated in this part of the experiment.

No physiological interpretation of this influence may at present be made. Psychologically, it appears that the response of subjects to flicker was learned during the training period and it is possible that they became confused with the temperature change and that it was necessary for them to learn a new criterion of vanishing flicker for new temperature components of the air pressure stimulus and that this new criterion differed from the former one. At temperature ranges from 35° to 45° C. some of the subjects reported that the air stimulus felt unusual in that it lacked any temperature component.

The results of Karstorf (4) show that vanishing flicker frequency for temperature stimuli without pressure (he used radiated stimuli) is much slower than any of the frequency values found in this experiment in which pressure was used. While the nature of the dependence of vanishing flicker upon temperature is not known the results show that temperature

appears to be an important factor. Several new problems for further experimentation are suggested by these data.

It was considered desirable to determine whether shifts in the physiological or psychological zero would influence vanishing flicker frequency when the factors of intensity, decremental time adaptation, and temperature of the air stimulus were held constant. Rather large and significant differences were found for two of the four subjects as a result of the two conditions of temperature adaptation used.

This finding suggests that the temperature influence might be related to the phenomenon of psychological zero; and that some of the subjects may be influenced more by acclimatization adaptation temperature than others, indicating that their discriminative criterion of flicker differs in this respect.

(Manuscript received April 10, 1936)

References

1. Allen, F. and Hollenberg, A., On the tactile sensory reflex, *Quart. J. Exper. Physiol.*, 1924, **14**, 352–378.
2. Allen, F. and Macdonald, P. A., The sensations of temperature, pain, and pressure, *Quart. J. Exper. Physiol.*, 1927, **16**, 321–332.
3. Bellows, R. M., An experimental isolation of some factors determining response to rhythmic cutaneous stimulation. I. Frequency, pressure, and time, J. Exper. Psychol., 1937, **20**.
4. Drury, M. B. and Dallenbach, K. M., The response of cold spots under successive stimulation, *Amer. J. Psychol.*, 1931, **43**, 490–496.
5. Earhart, H. and Dallenbach, K. M., The response of warm spots under successive stimulation, *Amer. J. Psychol.*, 1933, **45**, 724 f.
6. Karstorf, F., Ueber die Verschmelzung der Warmempfindung bei rhythmische erflogenden Reizen, *Zeitsch. Biol.*, 1920, **71**, 1–18.
7. Renshaw, S., The influence of solution temperature on the gustatory stimulus limens, *Psychol. Bull.*, 1933, **30**, 684–685.

THE RELATION BETWEEN AMOUNT OF MATERIAL AND DIFFICULTY OF PROBLEM SOLVING. I. MENTAL ADDITION AND SUBTRACTION

BY THOMAS W. COOK

Acadia University

Recent work by McGeoch and Oberschelp,[1] Scott and Henninger,[2] and Cook,[3] has indicated that with massed practice increase in number of elements conditions a disproportionate increase in the difficulty of simple problems. Moreover, the increase in difficulty seems greater than has been found by the majority of investigators using nonsense syllables or digits.

One plausible interpretation of these findings postulates a unique dependence of problem solving upon the range of mental grasp. For nonsense material, Thurstone's formulæ for the relation between complexity of material and difficulty contain a constant (a) designed to 'correct' for the subject's memory span.[4] Such a correction assumes that the memory span is a factor external to the kind of learning expressed by the equation, since it treats the individual span as a constant (additive) source of error that must be allowed for before the learning function can be expressed in a single dynamic pattern. Thus (a) may become zero without materially affecting the pattern, but if any other member of the equation is reduced to zero the whole equation equals zero.

For nonsense materials the facility with which Thurstone's formulæ fit the data is the only justification needed. But

[1] J. A. McGeoch and V. J. Oberschelp, The influence of length of problem and of transfer upon rational learning and its retention, *J. Genet. Psychol.*, 1930, **4**, 154–168.

[2] T. C. Scott and L. L. Henninger, The relation between length and difficulty in motor learning, J. EXPER. PSYCHOL., 1933, **16**, 657–678. The apparent exception in Warden and Hamilton, and Allison's data has been shown by Scott and Henninger to be due to the fact that those experimenters used a simple alternating pattern.

[3] T. W. Cook, Factors in whole and part learning a visually perceived maze (To appear in *J. Genet. Psychol.*).

[4] L. L. Thurstone, The relation between learning time and length of task, *Psychol. Rev.*, 1930, **37**, 44–53.

it does not follow that the formulæ are adequate to express the relation between complexity of *problem* and difficulty of solution. On the contrary, the deduction that problem solving takes place chiefly *within* the memory span (and is greatly increased in difficulty and perhaps even impossible without secondary aids when that span is exceeded) is sufficiently warranted by inspection of the data to justify adoption as a working hypothesis. The present investigation is the first phase of an attempt to subject this hypothesis to experimental test.

It is evident that any type of 'problem material' suitable for our purposes must admit increase in number of elements without qualitative change in the nature of the problem. Such a limitation considerably narrows the range of available problems: almost all puzzles, for example, being excluded. As a first essay in the field it was decided to use a graded series of problems in mental addition and subtraction.

Ten each of 2, 3, 4, 5, and 6 digit problems were constructed. The 2 digit problems comprised 5 addition problems and 5 subtraction problems, each 3 digit problem included one addition and one subtraction, the addition preceding the subtraction in 5 problems and succeeding the subtraction in 5 problems. In the 4, 5, and 6 digit problems the plus and minus signs occurred in random order, except that regular patterns (*e.g.*, $- + - + -$) were avoided and not more than two like signs appeared in succession. These exceptions to chance arrangement were designed to minimize the likelihood of a subject organizing the larger problems into patterns containing groups of like elements. In order to avoid the necessity of equating different amounts of previous practice upon one-place and two-place numbers, all numbers in the original problems, and all part answers or final answers were less than 10. Thus the first of the 6 digit problems was: $6 + 3 - 7 + 5 - 4 - 1$. The solution step by step yields the following values: $6 + 3 = 9 - 7 = 2 + 5 = 7 - 4 = 3 - 1 = 2$.

The apparatus consisted of sheets containing the 2, 3, 4, 5, and 6 digit problems typewritten in chance order, a screen to shut off S's view of the problem sheets and recording blanks, two stop clocks, and a metronome sounding at the rate of 100 strokes a minute. At the beginning of the experiment each subject [5] was asked to read the following instructions:

"You are to solve some problems in mental addition and subtraction. The problems will sometimes call for addition, sometimes for subtraction, and sometimes for a combination of addition and subtraction. (The number of digits will vary from 2 to 6, and the number of operations from 1 to 5.)

"Each problem will be read aloud by the experimenter, who will say 'ready' just before beginning to read the problem.

"Do *not* start to solve the problem until all the digits are read. The starting click of the time clock will be a signal that the experimenter has finished reading.

[5] Seniors and graduate students at Acadia University.

"Give your answer as soon as you get it, or tell E at once if you 'give up.' *Do not guess.* After each problem you will be questioned about your procedure. Be sure to tell E if you inadvertently began to compute before the time clock was started.

"Each problem will be repeated until you get the correct answer."

The instructions were then explained by E,[6] with emphasis upon the request to wait until the last digit had been read before starting to solve a given problem. As an additional precaution the first 10 problems were repeated at the end of the experiment and the results of the latter instead of the initial trials on these problems were used in computing the averages.

The experimenter now said 'ready,' and after a brief pause began to read the first problem at the rate of one digit (with sign) per metronome stroke. As soon as E pronounced the last digit clock A was started; when S proffered a solution (or 'gave up') clock A was stopped and clock B started. After 15 seconds, during which E said 'right' or 'wrong,' entered the results, and checked up on whether or not S had started to solve the problem before E had finished reading, the digits were repeated if S's answer had been incorrect, or a new problem was begun if a correct answer had been given.

A word should be said about the theoretical significance of one aspect of the above technique. It is clear that the relation of problem solving to the range of mental grasp can be satisfactorily tested only if it is possible to eliminate indirect perceptual aids *during* the solution of the problem. Ordinarily, of course, we do not depend upon 'ideating' or 'imagining' alone. We record symbols signifying the original items or the successive steps upon paper, and are then able to concentrate upon specific combinations of items without danger of loss of those features of the problem not under consideration at the moment. Such secondary perceptual assistance is in part deleted in the present instance by the auditory presentation, but there remained the possibility that S might start to solve a problem before E finished reading, and perform each operation as presented. To some extent this source of error was controlled by the rapid rate of reading (chosen for that purpose after preliminary trials), and by the random order of presentation of problems of different length. In the main, however, it depended upon the honesty and self-control of the subjects. Every effort was accordingly made to impress upon the latter the importance of this requirement.

Experimental Data.—The average time in seconds required by groups of three subjects to solve the 2, 3, 4, 5, and 6 digit

[6] Miss Jessie Harding, Laboratory Assistant in Psychology at Acadia University. The writer wishes to express his appreciation of Miss Harding's excellent work.

problems is presented in Table I,[7] together with the average time for all subjects, and the 'time per operation,' obtained by dividing the number of additions and subtractions per problem into the average time. The last named values show the degree to which the increase in difficulty exceeds the increase in complexity. An examination of the relation between the average time (for all subjects) and the number of operations reveals that for 3, 4, 5, and 6 digit problems the former can be expressed as a regular function of the latter: the difficulty (average time) increases as the cube of the complexity (number of operations). To demonstrate this we first multiply the average time for all subjects by 10, and get the values

TABLE I

AVERAGE TIME IN SECONDS

Subjects	Number of Digits				
	2	3	4	5	6
A, B, C	1.1	3.2	12.8	26.9	63.5
D, E, F	1.0	3.3	8.2	26.3	47.8
G, H, I	1.1	2.9	7.0	19.6	37.9
J, K, L	0.8	2.5	5.5	15.2	30.4
M, N, O	1.0	2.4	4.5	9.8	13.0
Average	1.0	2.8	7.6	19.6	38.5
Time per operation	1.0	1.4	2.5	4.9	7.7

10, 28, 76, 196, 386. The number of operations for each problem is then cubed and the resulting numbers: 1, 8, 27, 64, 125 (1, 2, 3, 4, 5), multiplied by the constant 3, obtaining the values 3, 24, 81, 192, 375—a close correspondence with average time except for the 2 digit (one operation) problems.

It must be noted, however, that the range of individual differences is wide, especially with the 6 digit problems. Table I shows that the three subjects making the best records (M, N, O) on the 6 digit problems require an average of 2.4, 4.5, 9.8, and 13.0 seconds for the 3, 4, 5, and 6 digit problems respectively. For the three least efficient subjects (A, B, C) the corresponding values are 3.2, 12.8, 26.9, and 63.5 seconds. That is, in the former group the difficulty increases approxi-

[7] The method of grouping only slightly distorts the picture presented by the individual records.

mately as the square of the complexity; in the latter group the difficulty increase is slightly greater than the cube of the complexity except from the 5 digit to the 6 digit problems (26.9 to 63.5 seconds), where it is about the fourth power. If we may accept 10 problems at each complexity as a fair sample of a subject's ability, the increase in difficulty is as much a function of the individual as the material. An even more striking increase in difficulty was exhibited in the record of one subject (P) whose results are not included in the general average because he was able to solve only 22 problems during the experimental period. His average speed on 3, 4, 5, and 6 digit problems was 11.5, 35.3, 43.9, and 190.4 seconds. Subject P's results point to the desirability of extending our work to include 7 digit and 8 digit problems. In the present instance it did not seem advisable to make the task too difficult for our volunteer subjects.

For trials Table II gives the results from 13 subjects

TABLE II

AVERAGE NUMBER OF TRIALS

Subjects	Number of Digits		
	4	5	6
Group I (B, C, D, E)	1.5	2.5	4.7
Group II (F, G, H, J, K)	1.3	2.5	3.6
Group III (I, L, M, N)	1.1	1.6	2.7
Average (including A and O)	1.3	2.2	3.6
Trials per operation	.43	.54	.72

(omitting A, O, and P) arranged in three groups. The grouping was determined by the fact that the members of a given group required almost exactly the same number of trials for problems of each length. Only the 4, 5, and 6 digit problems are shown in Table II, since all subjects except P solved 2 and 3 digit problems in one trial.[8] The corresponding values for the three a-typical subjects were: A, 1.9, 2.4, 4.0; O, 1.0, 1.3, 1.7; P, 2.5, 2.5, 7.8 trials. The averages of all

[8] Subject O required two trials on one 3 digit problem. Since O made no errors on any 4 digit problem this slip is probably of little significance.

subjects except P (1.3, 2.2, 3.6) indicate that for the 4, 5, and 6 digit problems the number of trials increases as the square of the number of operations. As with the time measure, however, individual differences are of first importance.

In general the data from both time and trials confirm the expectations derived from previous work. It is not intended, however, to enter into more detailed discussion of theoretical aspects of the data until further experiments have been performed.

Summary

Fifteen subjects solved ten each of 2, 3, 4, 5, and 6 digit problems in mental addition and subtraction.

The average time required to solve 2 and 3 digit problems differed only slightly, and neither required more than one trial per problem. For 3, 4, 5, and 6 digit problems the average time varied as the cube of the number of operations; for 4, 5, and 6 digit problems the average trials varied as the square of the number of operations.

The wide range of individual differences must be considered in estimating the significance of the data.

A COMPARISON OF CHOREIC WITH NORMAL CHILDREN ON THE BASIS OF SIMPLE REACTION TIMES TO VISUAL AND AUDITORY STIMULI

BY DOROTHY ALIENE BRADSHAW

Northwestern University

Introduction

The object of this study was to ascertain by extensive experimentation in the field of simple visual and auditory reaction times the validity of the observations often made by physicians to the effect that the voluntary responses of choreic children are fundamentally more rapid than those of so-called normal children. The form of chorea found in the subjects of this experiment belonged to the class termed 'chorea minor' by Osler (**12**) and 'common form' of chorea (Sydenham's) by Aldrich (**1**). This disorder recognized for the most part as an organic disease of the brain is characterized by widespread motor incoördinations and emotional instability. The child engages in many purposeless movements and all his voluntary movements, at first appearing quick and jerky, later develop to such an extent that apparently he is unable to perform any purposive movement. He becomes fretful, excitable, and irritable and appears to have difficulty in paying attention.

Whereas no studies have been made of the reaction times of choreic children, the relation of movement to speed of reaction has been the subject of much investigation. As shown by Feré, Cattell, Poffenberger, Gatewood, and others, the character of the reacting movement itself has a great deal to do with the time of reaction. Seashore (**15**) found a tendency for short reactions to be made by subjects with good voluntary movements, while Bagley (**2**) in a paper on the mental and motor ability of school children reported a general inverse relationship between motor ability and speed of re-

action. The reaction times upon which Bagley based his conclusions were reported by Miss Chapman (**4**) and included a study of choice and discrimination times as found by Jastrow's card sorting apparatus.

Investigations concerning the role of attention in reaction times have shown that the times of reaction are lessened as attention to the process is increased, and that the times are lengthened by distractions (**5**). When the attention of the subject is concentrated upon the movement, the resulting reaction has been found in most cases to be shorter than when the attention is concentrated upon the stimulus, although there is a tendency for this difference to disappear when the movement becomes habitual.

The reaction times of normal children were studied in Francis Galton's Anthropometric Laboratory (**11**, **13**), and later by Gilbert (**8**), Johnson (**10**), and Surber (**16**), but none of these investigators continued their experiments for a sufficient length of time to determine beyond a reasonable doubt whether or not definite reaction times for children could be established upon which times practice would have only a negligible effect, although it has long been recognized that practice is an important factor in the speed of reactions of adults and the results of Miss Johnson's study of the simple visual reaction times of children indicate that such times decrease with practice. Variations in the amounts of practice required by individual children and pronounced fluctuations in the average reaction times from day to day seem to the writer to indicate the importance and necessity of extensive practice before comparison of groups may be made.

Subjects

Of the twelve children who acted as subjects in the present study, six were choreic patients at the Children's Memorial Hospital in Chicago. These children, five of whom were girls and one a boy, ranged in age from seven to twelve years. No auditory or visual defects were noted by the examining physicians. The choreic conditions of the subjects varied widely, some patients manifesting very active symptoms while in other cases the disabilities were fairly mild. One patient was forced to remain in the hospital for a period of four months while another was released after thirty days. No attempt was made, however, to perform the experiments when each child was at approximately the same stage of recovery.

The six children in the control group were residents of the Illinois Children's Home and Aid Society of Evanston and were selected by the physician and superintendent

of that institution as being 'normal' both mentally and physically. Physical examinations showed that the children of this group had no visual or auditory defects except in the case of one boy whose slight visual defect was corrected by glasses. In no case was there any history of chorea. All these children—three boys and three girls—were doing well in school. Like the children of the experimental group, they ranged in age from seven to twelve years.

Apparatus and Procedure

Apparatus.—The apparatus used in this study was the standard Stoelting psychodometer with the 'downward pressure type' reaction key. The apparatus was the direct current model with power-driven tuning fork. A wooden screen—painted a very light grey—thirty-six inches long and eighteen inches high, placed in front of the apparatus from the position of the subject, concealed the experimenter and the equipment from the subject. Inserted in the screen was a piece of ground glass, two and one quarter inches in diameter, which allowed the child to see a green light furnished by the visual stimulus lamp but did not permit any view of the equipment.

Frequency of Experiments.—With both the choreic and the normal children the experiments were conducted six days a week. The choreic children were tested in the mornings and the normal children in the afternoons as the latter attended school and were not free until that time.

Testing Rooms.—The testing rooms at both institutions where the experiments were conducted were well lighted and well ventilated and the temperature was kept as constant as possible from day to day. The doors were closed during the experiments and through them could be heard practically no sounds of what went on outside the rooms which were located on floors where there was very little noise. The apparatus was so situated that the subject during the tests had his back toward the few furnishings. Only two people were in the room at one time—the subject and the examiner.

Stimuli.—Two forms of stimuli were used—light and sound. In each experiment eighty reactions were recorded. Of this number forty were reactions to the visual stimulus and forty were reactions to the auditory stimulus. To three of the choreic children and to three of the normal the visual stimulus was presented before the auditory and for the other six children the order was reversed but the order of presentation for each subject remained the same.

The visual stimulus was the flash of a green light provided by the visual stimulus lamp. The flash of light could be seen through the ground glass inserted in the screen. The light was approximately fifteen inches from the eyes of the subject and this distance was kept constant for all subjects by adjusting the height of the subject's chair and his distance from the screen.

The auditory stimulus was provided by a telegraph sounder fastened upon the base of the subject's equipment and concealed by the screen. The sounder was about twenty-four inches from the subject.

Method of Reacting.—The subject reacted to the stimuli by pressing lightly upon the reaction key.[1] During the experiments the choreic children were seated in wheel chairs and the normal children in arm chairs. A board six inches wide and approximately twenty-seven inches long was placed across the arms of the chair to serve as a support for the child's forearm and wrist as well as the reaction key. The child's elbow rested upon the arm of the wheel chair and his lower arm and wrist upon the board.

[1] In his study of the reaction times of children, Surber (16) found this type of key to be the most advantageous.

The index finger of his right hand rested lightly upon the key. All the children used a wrist and finger movement.

Instructions to the Subjects.—Instructions were repeated every day that the child keep his right arm in the assigned position and that he press the key as soon as he saw the light or heard the sound. At the beginning of the experiments the child was told just what to expect in the way of stimuli and five preliminary trials with each stimulus were given. In presenting the stimuli, the 'ready' signal was invariably given and the time interval between the signal and the presentation of the stimulus was varied between one and three seconds. By varying the interval within these limits it was possible to eliminate the occurrence of premature reactions in all but the first few experiments and the few that did appear were so obviously discordant that they could be eliminated at the time of experimentation so that a total of eighty reactions could be procured in every experiment.

Attitude of the Subjects.—The children were always perfectly at ease and enjoyed the tests and looked forward to the next day when there would be more 'games.' In spite of the seeming monotony of the work this interest was maintained throughout the entire series of experiments.

Results

Table I gives a brief summary of each subject's individual record and shows the arithmetic means and standard deviations of the reactions to each stimulus for each daily experiment listed. The reaction times reported are expressed in sigma and in every case represent the calculations upon a series of forty reactions to each of the stimuli employed. In Table II are shown the ultimate reaction times obtained for each subject and the number of experiments necessary in each case before the reaction times became stable.

Discussion of Results

Effect of Practice.—The records of all the subjects, both choreic and normal, show the almost inestimable importance of practice on simple reaction times to visual and auditory stimuli. With a sufficient amount of training under as uniform conditions as possible the reaction times of both choreic and normal children were found to reach ultimately a point beyond which it seemed practice would have a negligible effect—that is, the reaction times became stable. The amount of practice required before this limit was reached varied with the individual but the amount of practice required by the choreic children exceeded that required by the normal children and, furthermore, there was greater variation in the amount of

TABLE I

ARITHMETIC MEANS AND STANDARD DEVIATIONS OF REACTION TIMES TO EACH STIMULUS FOR FIRST, TENTH, TWENTIETH, TWENTY-FIFTH AND THIRTIETH EXPERIMENTS

Subj.	Physical Diagnosis	1st Experiment				10th Experiment				20th Experiment				25th Experiment				30th Experiment			
		Visual		Auditory		Visual		Auditory		Visual		Auditory		Visual		Auditory		Visual		Auditory	
		R.T. in Sigma	S. Dev.	R.T. in Sigma	S. Dev.	R.T. in Sigma	S. Dev.	R.T. in Sigma	S. Dev.	R.T. in Sigma	S. Dev.	R.T. in Sigma	S. Dev.	R.T. in Sigma	S. Dev.	R.T. in Sigma	S. Dev.	R.T. in Sigma	S. Dev.	R.T. in Sigma	S. Dev.
A	Chorea	229.1	61.7	221.2	51.4	176.1	34.0	167.2	38.1	184.6	25.5	200.8	29.2	165.7	18.8	158.5	23.6	137.4	4.6	134.8	6.1
B	"	187.7	27.9	214.7	37.4	157.3	20.2	160.7	29.5	177.6	20.2	145.7	16.0	163.7	26.9	175.1	25.6	134.5	2.0	130.1	3.2
C	"	181.0	25.7	165.5	27.6	190.5	27.0	182.0	25.2	129.5	4.0	131.6	10.9	129.2	3.1	122.1	2.5	129.2	3.5	122.3	2.4
D	"	175.4	32.1	176.7	30.9	179.1	22.8	144.6	18.3	140.8	12.3	136.6	12.2	129.0	3.5	122.8	2.5				
E	"	232.7	63.1	196.2	48.8	175.8	23.4	156.2	15.5	129.4	3.5	123.9	2.8	129.3	2.3	122.3	2.2				
F	"	182.3	28.5	190.2	26.6	156.4	16.2	148.3	16.1	130.0	4.0	128.8	7.0								
I	Normal	144.2	17.9	148.6	19.6	136.7	5.0	125.3	4.5	129.2	1.7	122.3	1.9								
II	"	145.0	13.7	138.7	12.8	139.6	5.9	137.5	11.1	129.1	2.1	122.1	1.5								
III	"	141.5	16.9	144.4	14.8	141.6	7.0	127.6	7.4	129.4	1.7	122.1	1.7								
IV	"	140.7	15.0	131.7	10.7	138.0	6.5	124.6	3.8	129.4	2.5	122.2	1.7								
V	"	142.9	10.1	137.2	8.2	132.4	4.7	124.2	3.8	129.0	2.2	122.0	1.6								
VI	"	168.3	16.4	151.6	14.9	140.4	5.1	140.6	5.7	129.2	1.9	122.0	1.2								

TABLE II

REACTION TIMES FOR EACH SUBJECT AFTER ESTABLISHMENT OF STABLE LEVEL OF PERFORMANCE

Subject	Sex	Age to Nearest Birthday	Group	Total No. of Exper.	No. of Exper. Before Auditory R.T. Became Stable	No. of Exper. Before Visual R.T. Became Stable	R.T. Last Experiment			
							Visual		Auditory	
							R.T. in Sigma	S. Dev.	R.T. in Sigma	S. Dev.
A	F.	8 yrs.	Chor.	37	33	32	127.23	2.15	122.43	2.06
B	F.	9 "	"	40	31	30	129.30	2.02	122.15	2.06
C	F.	11 "	"	34	24	18	129.28	2.85	122.43	2.35
D	F.	12 "	"	29	22	23	129.25	2.85	122.88	2.22
E	M.	7 "	"	27	20	18	129.10	2.00	122.08	1.77
F	F.	11 "	"	24	22	19	129.18	3.16	123.23	2.97
I	M.	7 "	Norm.	23	12	12	129.45	1.50	122.43	1.46
II	F.	9 "	"	23	14	13	129.15	1.75	122.10	1.36
III	F.	10 "	"	23	14	13	129.33	1.67	122.18	1.67
IV	F.	12 "	"	23	12	13	129.23	2.17	122.00	1.61
V	M.	9 "	"	23	11	10	129.45	2.20	122.03	1.04
VI	M.	11 "	"	23	14	14	129.03	1.72	122.03	1.21

practice among the choreic patients than among the members of the control group.

These results concerning the effects of practice are not comparable with those of other investigators since no one working with the reaction times of children has based his report on ultimate reaction times. In all previous studies only a few trials, comparatively speaking, have been given and the calculations based on the entire series. The results of Johnson's (**10**) study of children seem to indicate that the times of simple visual reaction decrease with practice but this investigator did not continue the experiments for a sufficient length of time to determine how rapid the reactions would become.

Variation with Age.—For the ages studied—seven to twelve years inclusive—neither the ultimate times of reaction nor the amount of practice required show any variation with age. Such a finding differs from the results reported by investigators of the simple reaction times and the times of discrimination and choice of children in that these workers with the possible exception of Johnson (**10**) have unanimously found a

decrease in speed of reaction with an increase in age. This same tendency was noticed in Johnson's study but no definite conclusions could be drawn concerning this factor because of the limited number of subjects and ages represented. The results of the present investigation would seem to indicate no age differences for the limited range of ages studied but it is possible some variation in reaction times with age might be found with a wider range of ages and a larger number of subjects.

Sex Differences.—Investigators of the reaction times of normal adults have generally discredited the importance of sex differences in this field and it is interesting to note that in the present study no variation with sex is apparent in the ultimate reaction times to either the visual or auditory stimulus nor in the amount of practice required. Such findings, however, are different from those of Gilbert (**8**) and Surber (**16**) both of whom noted sex differences in their respective studies of the reaction times of normal children. Using a visual stimulus, Gilbert found that boys react more quickly than girls while Surber, using an auditory stimulus, found that except for ages eight and nine, girls react more rapidly than boys.

Individual Differences.—Since none of the investigators of the reaction times of children continued their experiments for a sufficiently long period of time to outweigh the effects of practice, no definite conclusions can be drawn from their works concerning individual differences in this respect. Henmon and Wells (**9**) in their study with adults reported the presence of individual differences persisting after long continued practice. Whipple (**17**) and others, also working with normal adults, had previously reported the disappearance of individual differences with sufficient practice as determined by a mean variation of less than one-tenth the average time.

In the present investigation, the amount of practice required before the reaction times became stable varied with the individual children. No individual differences are found in the ultimate reaction times of the normal children. Among the choreic patients, the individual differences in the times of reaction to both the visual and auditory stimuli are, with one

possible exception, so slight as to be negligible. The ultimate auditory reaction time of four of the choreic children is 122 sigma and that of the other two children in this group approximately 123 sigma, the difference being fractional. Such small differences might be due to errors in technique. Furthermore, it is possible that with further practice the auditory reaction time of Subject F would become similar to that of the normal children so that the ultimate auditory reaction time of all the choreic children would be the same as that of the normal children. With the exception of Subject A the ultimate visual reaction time of the choreic children is the same as that of the normal children. In this case, the difference although slight is quite definite and consistent for that individual and it might be assumed that with more subjects such individual differences would be found also among the normal children.

Variation with Stimuli.—The results of the present study agree with those reported by most investigators of the reaction times of adults and with the findings of the Galton Anthropometric Laboratory in regard to the reaction times of children in that the times of simple reaction vary with the stimulus employed—the reactions to the auditory stimulus being more rapid than the reactions to the visual stimulus. With each individual child, choreic and normal, the auditory and visual reaction times became stable after approximately the same amount of practice, and no difference dependent upon order of presentation of stimuli was found in either the ultimate reaction times or the amount of practice required. The ultimate auditory reaction time for both groups of children as found in the present study (122 sigma) closely approximates the auditory reaction times of adults as found by Wundt, Cattell, and others. The ultimate visual reaction time (129 sigma) is shorter than the visual reaction times of adults as reported by most investigators but approximates that found by Johnson (**10**) in the cases of two children, eight and ten years of age. The difference, amounting to about seven sigma, between the times of visual and auditory reactions as found in the present investigation is practically identical with

that of eight sigma reported by Donders and others but is considerably less than that found by the workers of the Galton Laboratory and by investigators of the reaction times of adults. The significance of this smaller variation than is usually reported cannot be determined in the light of our present knowledge since the cause of a variation is not known but is only a matter of theory.

Effects of Chorea.—It should be noted that the attainment of the ultimate times of reaction does not depend upon or is not synonymous in time with the recovery of the choreic patients. This fact is strikingly shown in the case of Subject C who remained in bed in the hospital long after the experiments had been discontinued. During the entire series of thirty-four experiments the choreic symptoms were very active. The case of Subject B also clearly illustrates this point. This little girl remained in the hospital thirty-seven days after the completion of the experiments. On the day of the thirty-eighth experiment she was started on the nirvanol treatment and consequently the experiments were discontinued for thirteen days. Up to this time the choreic symptoms had been very pronounced but the ultimate times of reaction had been reached six or seven days before.

As has been previously stated, there is greater variability in the amount of practice required by the choreic children than by the normal children. This variability among the choreic subjects seems roughly proportional to the severity of the disease. The amount of practice required by Subjects A, B, and C exceeds that required by Subjects D, E, and F whose choreic symptoms were not as severe. However, Subject C, although exhibiting symptoms of approximately equal severity as those manifested by Subjects A and B, required considerably less practice.

Conclusions

1. With sufficient practice, the simple visual and auditory reaction times of both choreic and normal children become stable. Beyond this point, practice has little or no effect.

2. The amount of practice required before this limit is reached varies with the individual child.

3. The amount of practice required by choreic children exceeds that required by normal children.

4. There is considerable variation in the amount of practice required by choreic children. With normal children this variation seems much smaller.

5. The amount of practice required by choreic children is roughly proportional to the severity of the disease—the more severe cases requiring a greater amount of training.

6. The attainment of stable reaction times does not depend upon recovery from chorea.

7. Variations in the amount of practice required by individual children and fluctuations in the arithmetic means from day to day seem to indicate the futility of comparing groups of individual children on the basis of reaction times unless these times have become stable.

8. The reaction times of normal children do not become stable until approximately 500 reactions have been made.

9. With each individual child—choreic or normal—auditory and visual reaction times become stable after approximately the same number of trials.

10. The ultimate reaction times of choreic and normal children are practically identical. With more subjects to show the influence of individual differences, the records of both groups in all probability would be identical.

11. Individual differences are of little importance in the ultimate simple reaction times of children. Just one definite example was found in the choreic group and none in the control group.

12. No variation with sex is apparent in the ultimate reaction times to either the visual or auditory stimuli nor in the amount of training needed.

13. For the ages included—seven to twelve years inclusive—no variation with age in ultimate reaction times nor in the amount of practice required is apparent.

14. The ultimate visual reaction time of children is shorter than the visual reaction times of adults as reported by most investigators but closely approximates that found by Johnson in the cases of two children eight and ten years of age.

15. The ultimate auditory reaction time of children approximates the auditory reaction time of adults as found by Wundt, Cattell, and others.

16. With children, the ultimate time of reaction to a visual stimulus exceeds that to an auditory stimulus by about seven sigma.

17. This difference between the times of visual and auditory reactions is practically identical with that reported by Donders but is less than that found in the Galton Anthropometric Laboratory and by many investigators of the reaction times of adults.

(Manuscript received May 18, 1936)

References

1. Aldrich, C. J., Chorea: what is it?, *Arch. Ped.*, 1903, **20**, 827–837.
2. Bagley, W. C., On the correlation of mental and motor ability in school children, *Amer. J. Psych.*, 1901, **12**, 193–205.
3. Cattell, J. McK., On reaction time, *Mem. Nat. Acad. Sci.*, 1893, **7**.
4. Chapman, A., *The Development of Choice and Distinction*, Baccalaureate Thesis, Univ. of Wis., 1898.
5. Evans, J. E., The effect of distractions on reaction time with special reference to practice and the transfer of training, *Arch. Psych.*, 1916, **37**, pp. 106.
6. Feré, Ch., L'energie et la vitesse des mouvements voluntaires, *Rev. Phil.*, 1889, 38.
7. Gatewood, E. L., Individual differences in finger reactions, *Psychol. Monog.*, 1920, **28**, No. 4, 1–43.
8. Gilbert, J. A., Researches on the mental and physical development of school children, *Stud. Yale Psychol. Lab.*, 1894, **2**.
9. Henmon, V. A. C. and Wells, F. L., Concerning individual differences in reaction time, *Psychol. Rev.*, 1914, **21**, 153–156.
10. Johnson, B. J., Experimental study of motor abilities in the primary grades, *Johns Hopkins Stud. in Educ.*, 1917, **2**.
11. Koga, Y. and Morant, G. M., On the degree of association between reaction times in the case of different senses, *Biometrika*, **15**, 346.
12. Osler, W., *On chorea and choreiform affections*, London, H. K. Lewis.
13. Pearson, K., *Life, letters and labours of Francis Galton*. II, Cambridge, University Press, 1924.
14. Poffenberger, A. T., Reaction time to retinal stimulation, *Arch. Psychol.*, 1912, **23**.
15. Seashore, C. E., Motor ability, reaction-time, rhythm and time sense, *Univ. of Iowa Stud. Psychol.*, 1899, **2**, 64–84.
16. Surber, P., *Reaktionem auf Schallreize*, Zurich, Leeman, 1931.
17. Whipple, G. M., Reaction-time as a test of mental ability, *Amer. J. Psychol.*, 1904, **15**, 489–498.

THE RELATIVE EFFECTIVENESS OF EMPHASIS UPON RIGHT AND WRONG RESPONSES IN HUMAN MAZE LEARNING[1]

BY SIDNEY B. SILLECK, JR., AND CHAS. WESLEY LAPHA

Princeton University

The Problem

Investigators have failed to agree upon the effectiveness of emphasis upon right responses as compared with emphasis upon wrong responses in human maze learning. Tolman, Hall, and Bretnall (**7**) have reported that an electric shock and the sounding of a bell accompanying right responses is a more advantageous condition for learning than an electric shock and the sounding of a bell for wrong responses. Moreover, the same result was found when the bell alone was used to emphasize the response.

From a repetition and extension of this experiment Muenzinger (**4**) concluded that the superiority of the Bell-Right situation over the Bell-Wrong situation was not proved. Hulin and Katz (**2**) repeated that part of the Tolman-Hall-Bretnall experiment in which the electric shock was omitted and the bell alone used. They found a difference in favor of the Bell-Right situation and concluded that the evidence from the previous experiments also favored the Bell-Right situation.

To this conclusion Muenzinger has raised two objections: (*a*) the patterns learned by the Bell-Right and Bell-Wrong groups were not identical in all experiments and (*b*) the superiority of the Bell-Right group over the Bell-Wrong group failed to appear in one of his experiments. Earlier, Goodenough (**1**) criticized the interpretation of Tolman, Hall, and Bretnall in part because the differences found between the Bell-Right and Bell-Wrong groups could be due to differences in the learning abilities of the two groups rather than to the experimental conditions.

[1] From the psychological laboratory of Princeton University.

The present experiment, therefore, is an attempt to take into account the two possibilities of difference in pattern and difference in learning ability in obtaining further data on the relative effectiveness of emphasis upon right and wrong responses in human maze learning.

METHOD

Two mazes were used in this experiment. The one was a duplicate of the Tolman Punchboard Maze (7) [the identical maze used by Hulin and Katz (2)]. The other was a duplicate of the stylus-maze used by T. C. Scott (6) and called by him 'maze C' (the design originating with C. J. Warden). The punchboard maze contained thirty pairs of holes, irregularly placed in a 3′ by 3′ board. One hole of each pair had an electric connection to the bell-stimulus, but the particular holes so wired were the opposite holes to the ones which Tolman connected with the bell-stimulus in his main experiments. The stylus-maze contained ten cul-de-sacs. The subject found his way through the grooves by means of a wooden stylus. A screen prevented the subjects from seeing the stylus-maze.

Two groups of subjects, each containing 25 undergraduates, were tested; one group (Bell-Right group) learned the punchboard maze under instructions to place the stylus only in the holes that sounded the bell. The other group (Bell-Wrong group) learned the punchboard maze under instructions to place the stylus only in the holes that did not sound the bell. The procedure adopted by Tolman, Hall, and Bretnall was followed for the learning of the punchboard maze. Each subject was given three practice trials in which he placed the stylus in both holes of each of the thirty pairs. Then he was told whether to place the stylus in the ringing or non-ringing holes and was instructed to correct an error immediately. The order of holes to be followed was given by numbers on the maze. The criterion of learning was one errorless trial, but subjects who did not meet this criterion within twenty trials were stopped at that point.

The Bell-Wrong group learned the sequence of holes called by Muenzinger Pattern A and the Bell-Right group the opposite pattern, or Pattern B. In other words, the experimenters made no changes in the electrical wiring of the punchboard for the two groups. The learning of Pattern A by the Bell-Wrong group and Pattern B by the Bell-Right group is a reversal of the situation in the main experiments of Tolman, Hall, and Bretnall and of Hulin and Katz. In both of these experiments Pattern A was learned by the Bell-Right group and Pattern B by the Bell-Wrong group. Inasmuch as the superiority of the Bell-Right groups in these experiments may have been the result of learning Pattern A, in the present experiment the Bell-Right group was made to learn Pattern B.

The stylus-maze was used to compare the relative maze-learning abilities of the two groups. Approximately half of the Bell-Right group learned the stylus-maze prior to the punchboard maze; the remaining Bell-Right subjects learned the punchboard first. Similarly the possibility of transfer and practice effects in the learning of the two mazes was controlled for the Bell-Wrong group. The criterion of learning for the stylus-maze was one errorless trial, but no subject was given more than twenty trials.

RESULTS

Table I shows the total number of errors and the total number of trials for the individual subjects of the Bell-Right

and Bell-Wrong groups on the punchboard maze. All the Bell-Right subjects learned the maze in twenty trials, while nine of the Bell-Wrong subjects failed to meet this criterion. The average number of trials for the Bell-Right group was 11.7; the average number of trials for the Bell-Wrong group

TABLE I

SHOWING TOTAL ERRORS AND TRIALS FOR EACH SUBJECT IN THE TWO GROUPS ON THE PUNCHBOARD MAZE

Bell-Right			Bell-Wrong		
Subject	Total No. Errors	Total No. Trials	Subject	Total No. Errors	Total No. Trials
Fa	0	1	Sto	23	9
Ge	4	4	Mco	31	6
Wo	9	4	Do	35	11
Mo	24	7	Ba	53	11
Go	24	7	Ra	55	12
Hi	26	8	Sw	78	13
Mc	29	7	Fal	78	13
Va	39	8	Wh	79	11
St	49	15	Br	85	15
Dz	51	12	Coo	88	12
Le	52	9	Cor	100	14
Ta	55	9	Ha	103	16
Fo	61	13	Ac	112	16
Li	69	16	Ma	112	17
Ag	72	12	Wal	112	16
Tay	79	16	Cl *	135	20
Fi	80	12	Kl	150	20
Gr	88	15	Sm *	154	20
Wa	93	16	Fr *	155	20
Wi	95	18	Mcc *	181	20
La	96	17	Ban *	182	20
Lei	112	17	Law *	185	20
Ro	114	15	Mon *	188	20
Co	145	18	He *	201	20
Ot	152	17	Pa *	213	20
Sum	1618	293	Sum	2888	392
Average	64.7	11.7	Average	115.5	15.7

* Indicates not learned within 20 trials.

15.7. This difference of 4.0 trials is 4.61 times its probable error. The average number of errors for the Bell-Wrong group was 115.5 with a range of from 23 to 213 errors for the individual subjects. The average error for the Bell-Right group was 64.7 with a range of from 0 to 152 for the individual

subjects. The difference in average error in favor of the Bell-Right group is 50.8 and is 5.5 times its probable error.

The maze-learning abilities of the Bell-Right and Bell-Wrong groups were compared by the check experiment which utilized the stylus-maze. Table II presents the total number

TABLE II

SHOWING TOTAL ERRORS FOR EACH SUBJECT ON THE STYLUS-MAZE

Bell-Right Group		Bell-Wrong Group	
Subject	Total No. Errors	Subject	Total No. Errors
Va	6	Wal	8
Li	9	Fal	17
Tay	12	Mco	23
Ro	16	Sto	24
Go	16	Br	25
Wo	17	Cl	25
Ge	17	He	28
Wa	24	Ra	33
Wi	32	Ha	39
Fo	32	Mon	42
Le	33	Ac	42
Fa	33	Sm	44
St	39	Law	46
Mo	41	Wh	49
Ot	43	Coo	52
Mc *	48	Ma *	57
La *	49	Sw *	59
Ta	50	Con *	61
Fi	53	Do *	69
Lei *	56	Kl *	69
Co	58	Fr *	71
Hi	59	Mcc *	72
Dz *	60	Ba *	78
Ag *	64	Pa *	83
Gr *	75	Ban *	96
Sum	942	Sum	1212
Average	37.7	Average	48.5

* Indicates not learned within 20 trials.

of errors for the individual subjects in these two groups on the stylus-maze. The average error on the stylus-maze for the Bell-Wrong group was 48.5; the average error for the Bell-Right group was 37.7. This difference of 10.8 is 2.73 times its probable error. Thus, while the subjects for the Bell-Right and Bell-Wrong groups were selected at random, the two groups were not equal in maze-learning ability as tested by performance on the stylus-maze.

The superiority of the Bell-Right group over the Bell-Wrong group on the punchboard maze, however, cannot be attributed solely to the chance presence of better maze learners in the Bell-Right group. The difference between these groups on the punchboard maze is much greater than the difference on the stylus-maze. On the punchboard maze the difference is 5.5 times its probable error; on the stylus-maze 2.73 times its probable error. It, therefore, appears that the experimental conditions under which the two groups learned the punchboard (namely the sounding of a bell for the right responses for one group and the sounding of a bell for the wrong responses for the other group) have contributed to the superiority of the one group.

The product-moment correlation between errors on the punchboard maze and on the stylus-maze for the fifty subjects was .38 ± .08. The stylus-maze in itself, therefore, is not an adequate check upon the learning ability on the punchboard maze.

Table III summarizes the experiments on this problem to

TABLE III

	Average No. of Errors		Diff.	C.R.
	Bell-Right	Bell-Wrong		
Tolman, Hall, and Bretnall	103.81	117.22	13.41	1.95
Muenzinger (Pattern A)	116.1	145.2	29.1	3.3
Muenzinger (Pattern B)	125.5	124.6	−.9	.2
Hulin and Katz	58.2	76.0	17.8	2.3
Silleck and Lapha	64.7	115.5	50.8	5.5

date. Of the five experiments, four indicate a superiority of the Bell-Right situation over the Bell-Wrong situation and one shows practically no difference between the two situations.

It is possible, of course, that other factors besides the Bell-Right and Bell-Wrong conditions of learning account for these results. One factor may be the failure to equate the maze-learning abilities of the two experimental groups. In the present experiment the importance of this factor was established by comparing the groups on a stylus-maze. Since the

stylus-maze is learned without visual cues, it is conceivable that a maze more comparable to the punchboard maze might have demonstrated even greater differences in the learning abilities of the two groups. In spite of this conjecture, however, the fact remains that the difference in maze-learning ability demonstrated in the present experiment was not great enough to account for the difference in the Bell-Right and Bell-Wrong situations. Moreover, it is doubtful that the chance presence of better maze-learners would have occurred in sufficient numbers in the Bell-Right groups in all of the four experiments to explain the superiority of these groups.

A second factor which has been emphasized by Muenzinger (**5**) is the difference in patterns of holes learned by the two groups. In the experiments of Tolman, Hall, and Bretnall and Hulin and Katz, the Bell-Right group learned Pattern A and the Bell-Wrong group Pattern B. Hence the obtained differences in favor of the Bell-Right groups may mean that Pattern A is easier to learn than Pattern B. In the present experiment, however, the Bell-Right group learned Pattern B and the Bell-Wrong group Pattern A, and still the Bell-Right group made fewer errors than the Bell-Wrong group. Moreover, Muenzinger's experiments (**4**) in which the pattern was controlled by giving both groups the same pattern showed the Bell-Right situation to be superior by a difference 3.3 times its probable error in the one experiment and the Bell-Wrong situation superior by a difference only 0.2 times its probable error in the other experiment. And Muenzinger himself concludes that no reliable difference has been shown between Pattern A and Pattern B.

Since neither differences in maze-learning ability nor differences in pattern satisfactorily account for consistent differences in favor of the Bell-Right situation, the present writers interpret their data as a corroboration of the conclusion of Tolman, Hall, and Bretnall and Hulin and Katz that the sounding of a bell accompanying right responses is a more effective condition for learning than sounding of a bell accompanying wrong responses.

References

1. Goodenough, F. L., A note on Tolman's 'disproof' of Thorndike's law of effect, J. Exper. Psychol., 1933, **16**, 459–462.
2. Hulin, W. S., and Katz, D., A comparison of emphasis upon right and wrong responses in maze learning, J. Exper. Psychol., 1935, **18**, 638–642.
3. Hulin, W. S., and Katz, D., A comparison of the bell signal for right and wrong responses: A reply, J. Exper. Psychol., 1936, **19**, 117–118.
4. Muenzinger, K. F., Motivation in learning. II. The function of electric shock for right and wrong responses in human learning, J. Exper. Psychol., 1934, **17**, 439–448.
5. Muenzinger, K. F., A note on the relative efficacy of a bell signal for right and wrong responses in maze learning, J. Exper. Psychol., 1936, **19**, 116–117.
6. Scott, T. C., The retention and recognition of patterns in maze learning, J. Exper. Psychol., 1930, **13**, 164–207.
7. Tolman, E. C., Hall, C. S., and Bretnall, E. P., A disproof of the law of effect and a substitution of the law of emphasis, motivation, and disruption, J. Exper. Psychol., 1932, **15**, 601–614.
8. Warden, C. J., Primacy and recency as factors in cul-de-sac elimination in a stylus maze, J. Exper. Psychol., 1924, **7**, 98–116.